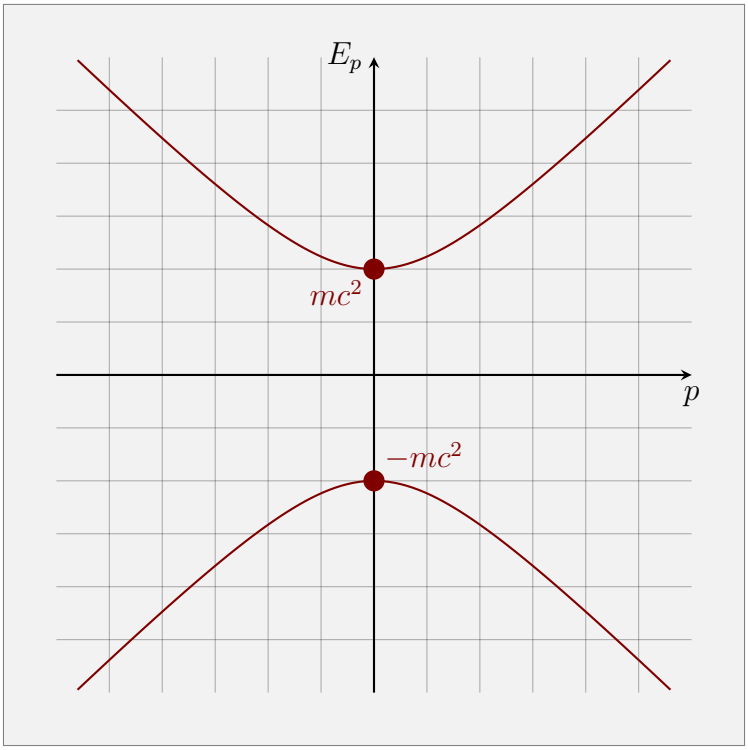
Advanced quantum mechanics

An introduction



Dominik Tschimmel

This page intentionally left blank.

Preface

Introductory quantum mechanics courses usually focus on one non-relativistic particle and/or on one space dimension. After the essentials of quantum mechanics are understood, the natural question is, what happens upon removing the restrictions. This is, where advanced quantum mechanics enters.

Fixing the issues of the Dirac theory, developed in advanced quantum mechanics, leads to a many particle theory. The description of the relativistic quantum many body system leads to objects, called Field operators. This is the onset of quantum field theory. The concepts of canonical quantization and functional integrals are introduced as outlook on qft.

Some parts of these notes were created, at the time I was learning the corresponding concepts. So be aware, that there may not only be the usual typos, but possibly wrong statements. In that sense, read with caution. However, nothing presented here is new, and usually well covered in textbooks.

This page intentionally left blank.

Contents

Part I.

Advanced quantum mechanics

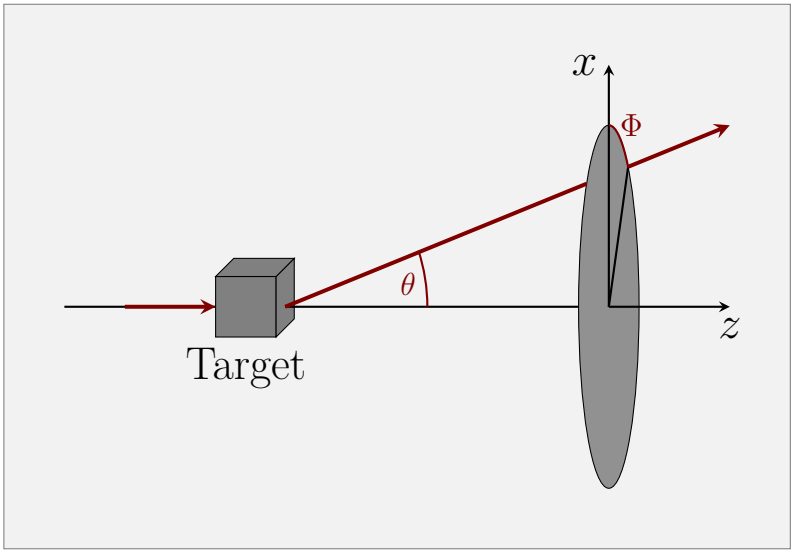
As the name suggests, the topics presented here are usually part of a followup course to introductory quantum mechanics. Introductory quantum mechanics focuses on non-relativistic single particles and scattering in one dimension. In this part, scattering in generalized to three dimensions. Also, a many particle theory, that goes by the name *second quantization* is developed. Finally, the attempt of a relativistic generalization is made by introducing the Dirac theory. This part concludes by combining second quantization with the Dirac theory, to solve the problems that arise from the latter alone. This leads in a natural way to quantum field theory.

1
Scattering theory

In modern physics, from high energy physics to solid state physics, lots of experiments are essentially scattering experiments. In that regard, the quantum mechanical description of scattering processes is of great importance to understand those experiments completely. This chapter is completely based on [Zir10], with some aspects from [Alt12].

1.1. Differential cross section

We consider a beam of particles that hits a target. Choosing the coordinates such that the initial beam moves along the z-axis, the setup can be depicted as follows:



The (probability) current density of the incoming particles is $j = j_0[dx \wedge dy, Or]$. The coefficient j_0 has the dimension

$$[j_0] = \frac{\text{number of particles}}{\text{time} \cdot \text{area}} .$$

After the scattering, the scattered particles move radially away from the scattering center, if observed from a large distance. Thus the current density is best described in spherical coordinates: $J = J_0[\sin(\theta)d\theta \wedge d\phi, Or]$. The coefficient here has the dimension

$$[J_0] = \frac{\text{number of particles}}{\text{time}} .$$

The current density should have the dimension $^{\text{particles}}/_{\text{time}}$. Since the 2-form $dx \wedge dy$ has the dimension length^2, there has to be an addition $^1/_{\text{area}}$ in the dimension of j_0. Notice that some authors use an additional $^1/_{\text{solid angle}}$ in the dimension of J_0. However this dimension is dimensionless in fact, though experimental physicists like to give it a unit, so we will omit it.

Definition 1.1.1.
The **differential cross section** $\frac{d\sigma}{d\Omega}$ is defined by

$$\frac{d\sigma}{d\Omega} = \frac{J_0}{j_0} \ .$$

It should be emphasized here, that the differential cross section is no derivative by any means. Mathematically it is a mere notation of the quantity. From the dimension of the coefficients j_0 and J_0 it follows that the differential cross section has the dimension *area*.

Definition 1.1.2.
The **total cross section** σ_{tot} is the differential cross section, integrated over the full solid angle:

$$\sigma_{tot} \coloneqq \int_{S^2} \frac{d\sigma}{d\Omega}\, d\Omega = \frac{1}{j_0} \int_{S^2} J \ .$$

1.2. Lippmann-Schwinger equation

So far, quantum mechanics has not entered in the discussion of scattering. In fact, the previous definitions hold for classical scattering as well.

1.2.1. Representation independent

We consider a time-independent Hamilton operator $\widehat{H}$, that can be written as sum $\widehat{H} = \widehat{H}_0 + \widehat{V}$ of a free (or known) Hamiltonian $\widehat{H}_0$ and a potential/perturbation $\widehat{V}$. Suppose $|\psi_0\rangle$ is a solution of the free Hamiltonian $\widehat{H}_0|\psi_0\rangle = E|\psi_0\rangle$. We are looking for a solution of the full Hamiltonian, with the same energy:

$$\widehat{H}|\psi\rangle = \widehat{H}_0|\psi\rangle + \widehat{V}|\psi\rangle = E|\psi\rangle \ .$$

Theorem 1.2.1.
Let $\widehat{G}_0 = (E + i\varepsilon - \widehat{H}_0)^{-1}$ be the inverse operator. If $|\psi\rangle$ is a solution of the equation

$$|\psi\rangle = |\psi_0\rangle + \widehat{G}_0\widehat{V}|\psi\rangle \qquad \Leftrightarrow \qquad |\psi_0\rangle = (\mathbb{1} - \widehat{G}_0\widehat{V})|\psi\rangle$$

then $|\psi\rangle$ solves $\widehat{H}|\psi\rangle = E|\psi\rangle$ for $\varepsilon \searrow 0$, if the full Hamiltonian $\widehat{H}$ has eigen states $|x_n\rangle$, that define a Hilbert basis.

Proof 1.2.2.
Rewriting of the equation yields:

$$|\psi_0\rangle = \widehat{G}_0 \left(\widehat{G}_0^{-1} - \widehat{V}\right)|\psi\rangle \ .$$

Define $\widehat{G} = (E + i\varepsilon - \widehat{H})^{-1}$, then from $\widehat{G}^{-1} = E + i\varepsilon - \widehat{H}$ and $\widehat{G}_0^{-1} = E + i\varepsilon - \widehat{H}_0$, it follows that

$$\widehat{G}^{-1} = \widehat{G}_0^{-1} - \widehat{V} \qquad \text{and thus} \qquad |\psi_0\rangle = \widehat{G}_0\widehat{G}^{-1}|\psi\rangle \ .$$

Left multiplication of the inverse operators results in:

$$|\psi\rangle = \widehat{G}\widehat{G}_0^{-1}|\psi_0\rangle = \widehat{G}\left(E + i\varepsilon - \widehat{H}_0\right)|\psi_0\rangle = i\varepsilon\widehat{G}|\psi_0\rangle \ .$$

Hence

$$\left(E - \widehat{H}\right)|\psi\rangle = i\varepsilon\left(E - \widehat{H}\right)\widehat{G}|\psi_0\rangle \ .$$

It remains to show, that $\left(E - \widehat{H}\right)\widehat{G}(\varepsilon)$ is bounded, such that for $\varepsilon \to 0$, the claim follows:

$$0 = \left(E - \widehat{H}\right)|\psi\rangle \qquad \Rightarrow \qquad \widehat{H}|\psi\rangle = E|\psi\rangle \ .$$

By assumption there is an eigen Hilbert basis $\{|x_n\rangle\}$, such that

$$\widehat{H}|x_n\rangle = E_n|x_n\rangle \qquad \Rightarrow \qquad \widehat{G}^{-1}|x_n\rangle = (E + i\varepsilon - E_n)|x_n\rangle \ .$$

This means, that $\{|x_n\rangle\}$ is a Hilbert basis of $\widehat{G}$:

$$\widehat{G}^{-1}|x_n\rangle = (E + i\varepsilon - E_n)|x_n\rangle \qquad \Leftrightarrow \qquad \widehat{G}|x_n\rangle = \frac{1}{E - E_n + i\varepsilon}|x_n\rangle \ .$$

In the following we will use the matrix representation w.r.t. the eigen basis $\{|x_n\rangle\}$. This does not need $\widehat{G}$ to be bounded (see remark A.1.8 for this):

$$\begin{aligned}
\langle\psi|(E - \widehat{H})\widehat{G}|\psi\rangle &= \sum_{m,n}\langle\psi|x_m\rangle\langle x_m|(E - \widehat{H})\widehat{G}|x_n\rangle\langle x_n|\psi\rangle \\
&= \sum_{m,n}\langle\psi|x_m\rangle\langle x_n|\psi\rangle\frac{E - E_n}{E - E_n + i\varepsilon}\langle x_m|x_n\rangle \\
&= \sum_{m,n}\langle\psi|x_m\rangle\langle x_n|\psi\rangle\frac{E - E_n}{E - E_n + i\varepsilon}\delta_{mn} \\
&= \sum_{n}|\langle x_n|\psi\rangle|^2\frac{E - E_n}{E - E_n + i\varepsilon} \ .
\end{aligned}$$

Boundedness means that $\|(E - \widehat{H})\widehat{G}\|_{Op} < \infty$. Lemma A.2.5 allows to restrict to $|\psi\rangle$ with $\||\psi\rangle\| \leq 1$. It follows that $\sum_n |\langle x_n|\psi\rangle|^2$ holds, and it remain to show $\left|\frac{E-E_n}{E-E_n+i\varepsilon}\right| \leq 1$ or equivalently $\left|\frac{E-E_n}{E-E_n+i\varepsilon}\right|^2 \leq 1$. For convenience we write $\chi_n = E - E_n$ in the following:

$$\begin{aligned}
\frac{\chi_n}{\chi_n + i\varepsilon} &= \frac{\chi_n(\chi_n - i\varepsilon)}{\chi_n^2 + \varepsilon^2} = \frac{\chi_n^2}{\chi_n^2 + \varepsilon^2} - i\frac{\chi_n\varepsilon}{\chi_n^2 + \varepsilon^2} \\
\Rightarrow \qquad \left|\frac{\chi_n}{\chi_n + i\varepsilon}\right|^2 &= \frac{\chi_n^4 + \chi_n^2\varepsilon^2}{(\chi_n^2 + \varepsilon^2)^2} \leq 1 \ .
\end{aligned}$$

Thus we have shown:

$$\begin{aligned}
&|\langle\psi|(E - \widehat{H})\widehat{G}|\psi\rangle|^2 \leq 1 \quad \forall\ |\psi\rangle \text{ with } \||\psi\rangle\| \leq 1 \\
\Rightarrow \qquad &\|(E - \widehat{H})\widehat{G}\|_{Op} \leq 1 \quad \forall\varepsilon \ .
\end{aligned}$$

□

Definition 1.2.3.
The equation $|\psi\rangle = |\psi_0\rangle + \widehat{G}_0\widehat{V}|\psi\rangle$ is called **Lippmann-Schwinger equation**. Here it is given in the general form without choice of representation.

1.2.2. Position representation

The Lippmann-Schwinger equation that we have used so far may not be the version that can be found in all textbooks. Its more common form is the equation that results using the position representation of quantum mechanics.

Let (as usual) the operators $\widehat{H}_0$ and $\widehat{V}$ be

$$\widehat{H}_0 = -\frac{\hbar^2}{2m}\nabla^2 \qquad \text{and} \qquad \widehat{V} = V(\vec{x}) \ .$$

Theorem 1.2.4.
The operator $\widehat{G}_0$ can be expressed as integral operator in the position representation

$$\langle \vec{x}|\widehat{G}_0|\varphi\rangle = \int_{\mathbb{R}^3} dy^3 \ \langle \vec{x}|\widehat{G}_0|\vec{y}\rangle\langle \vec{y}|\varphi\rangle \ ,$$

with Green's function

$$\langle \vec{x}|\widehat{G}_0|\vec{y}\rangle = -\frac{1}{4\pi}\frac{e^{ik|\vec{x}-\vec{y}|}}{|\vec{x}-\vec{y}|}\frac{2m}{\hbar^2} \ .$$

Proof 1.2.5.
Fundamental solution of the Laplace operator ∇^2 in $\mathbb{R}^3$:
We consider the function $f(\vec{x}) = \frac{1}{|\vec{x}|}$. With $\nabla^2 = \star d \star d$ for functions in $\mathbb{R}^3$ we calculate:

$$\begin{aligned}
(\nabla^2 T_f)[\varphi] &= \int_{\mathbb{R}^3\setminus\{0\}} f\nabla^2\varphi \ dx^3 = \int_{\mathbb{R}^3\setminus\{0\}} f\cdot \star d\star d\varphi \ dx^3 \\
&= \int_{\mathbb{R}^3\setminus\{0\}} f d\star d\varphi = \underbrace{\lim_{\varepsilon\to 0}\int_{\partial(\mathbb{R}^3\setminus B_\varepsilon(0))} f\star d\varphi}_{=0} - \int_{\mathbb{R}^3\setminus\{0\}} df\wedge\star d\varphi \\
&= -\int_{\mathbb{R}^3\setminus\{0\}} d\varphi\wedge\star df = -\lim_{\varepsilon\to 0}\int_{\partial(\mathbb{R}^3\setminus B_\varepsilon(0))} \varphi\wedge\star df + \int_{\mathbb{R}^3\setminus\{0\}} \varphi\cdot d\star df \ .
\end{aligned}$$

Since $f(\vec{x}) = \frac{1}{|\vec{x}|}$ is harmonic on $\mathbb{R}^3\setminus\{0\}$, i.e. $\nabla^2 f = 0$ and thus $d\star df = 0$, it follows that:

$$\begin{aligned}
(\nabla^2 T_f)[\varphi] &= -\lim_{\varepsilon\to 0}\int_{\partial(\mathbb{R}^3\setminus B_\varepsilon(0))} \varphi\wedge\star df = \lim_{\varepsilon\to 0}\int_{\partial(\mathbb{R}^3\setminus B_\varepsilon(0))} \varphi \ d\Omega \\
&= -\lim_{\varepsilon\to 0}\int_{\partial B_\varepsilon(0)} \varphi \ d\Omega = -4\pi\varphi(0) \ .
\end{aligned}$$

Hence the fundamental solution for the Laplace operator in $\mathbb{R}^3$ is

$$T_h \qquad \text{mit} \qquad h(\vec{x}) = -\frac{1}{4\pi|\vec{x}|} \ .$$

Fundamental solution for $E - \widehat{H}_0$:
Because of rotation invariance, we can simplify calculations using spherical coordinates. The Laplace operator becomes $\nabla^2 = \frac{1}{r}\partial_r^2 r + \Delta_{\phi,\theta}$, where $\Delta_{\phi,\theta}$ denotes the angular components, that vanish anyway because of rotation invariance.

The function $u(r) = \frac{2m}{4\pi\hbar}e^{ikr}$ is smooth, such that $u \cdot T_h = T_{u\cdot h}$ is well defined. Choose $g(r) = -u(r) \cdot h(r)$, then:

$$g(r) = \frac{-2m}{4\pi\hbar^2}\frac{e^{ikr}}{r} = \frac{-2m}{4\pi\hbar^2}\left(\frac{1}{r} - \frac{1-e^{ikr}}{r}\right) \ .$$

With $\nabla^2 = \frac{1}{r}\partial_r^2 r$ it follows that

$$\begin{aligned}\nabla^2\left(\frac{1-e^{ik|\vec{x}|}}{|\vec{x}|}\right) &= \frac{1}{r}\partial_r^2 r \cdot \left(\frac{1-e^{ikr}}{r}\right) + \Delta_{\phi,\theta}\left(\frac{1-e^{ikr}}{r}\right) \\ &= \frac{k^2 e^{ikr}}{r} = \frac{k^2 e^{ik|\vec{x}|}}{|\vec{x}|} \ .\end{aligned}$$

For the regular distribution T_g that means:

$$\nabla^2 T_g = \frac{-2m}{4\pi\hbar^2}\left(\nabla^2 T_{\frac{1}{r}} - \nabla^2 T_{\left(\frac{1-e^{ikr}}{r}\right)}\right) = \frac{2m}{\hbar^2}\delta(\vec{x}) + k^2 T_g \ .$$

The energy expressed in terms of the wave number k is $E = \frac{\hbar^2 k^2}{2m}$. Thus T_g is the fundamental solution for $E - \widehat{H}_0$:

$$\left(E + \frac{\hbar^2}{2m}\nabla^2\right) T_g = \delta(\vec{x}) + \left(-\frac{\hbar^2 k^2}{2m} + E\right) T_g = \delta(\vec{x}) \ .$$

Green's function for $E - \widehat{H}_0$:
Using corollary F.2.21 we get the Green's function

$$\langle\vec{x}|\widehat{G}_0|\vec{y}\rangle \equiv G(\vec{x},\vec{y}) = g(\vec{x}-\vec{y}) = -\frac{1}{4\pi}\frac{e^{ik|\vec{x}-\vec{y}|}}{|\vec{x}-\vec{y}|}\frac{2m}{\hbar^2} \ .$$

□

To be precise, we understand the Green's function integral operator as inverse of the corresponding differential operator, as it defines the (weak) solution of the differential equation by action on the inhomogeneity (see theorem F.2.15).

We finally see, how the **Lippmann-Schwinger equation** looks in the position representation:

$$\boxed{\psi(\vec{x}) = \psi_0(\vec{x}) - \frac{2m}{4\pi\hbar^2}\int_{\mathbb{R}^3}\frac{e^{ik|\vec{x}-\vec{y}|}}{|\vec{x}-\vec{y}|}V(\vec{y})\psi(\vec{y})\ dy^3}$$

1.3. Scattering amplitude

Revisiting the scattering setup from the beginning of this chapter in a quantum mechanical point of view means to investigate the *full wave function* $\psi(p)$ at a point p that is far away from the scattering center σ.

Assuming a scattering potential that has compact support (rapidly decreasing at least), the full wave function can be approximated by

$$\psi(p) \simeq \psi_0(p) + \frac{e^{ikr_\sigma(p)}}{r_\sigma(p)} f(\vec{k}', k, \psi) \ ,$$

with distance function $r_\sigma(p) = d(p, \sigma)$. The term $f(\vec{k}', k, \psi)$ is called **scattering amplitude**.

Although we have chosen the coordinate independent formulation as starting point, the calculations from [Alt12] carry over without changes. First, we write the Lippmann-Schwinger equation in the coordinate free way:

$$\psi(p) = \psi_0(p) - \frac{2m}{4\pi\hbar^2} \int\limits_{E_3} \frac{e^{ikr_p}}{r_p} V\psi \ dV$$

To obtain the desired approximation, we need an approximation for $r_p(q)$. As E_3 is an affine space, there are vectors $\vec{x}$ and $\vec{y}$, such that $p = \sigma + \vec{x}(p)$ and $q = \sigma + \vec{y}(q)$. It follows that

$$r_p(q) = |\vec{x} - \vec{y}| = \sqrt{|\vec{x}|^2 + |\vec{y}|^2 - 2\vec{x}\cdot\vec{y}} = |\vec{x}| \sqrt{1 + \frac{|\vec{y}|^2}{|\vec{x}|^2} - 2\frac{\vec{x}\cdot\vec{y}}{|\vec{x}|^2}} \ .$$

By construction we have $|\vec{x}| \gg |\vec{y}|$, allowing to use the following approximation

$$1 + \frac{|\vec{y}^2|}{|\vec{x}|^2} - 2\frac{\vec{x}\cdot\vec{y}}{|\vec{x}|^2} \simeq 1 - 2\frac{\vec{x}\cdot\vec{y}}{|\vec{x}|^2} \ .$$

The Taylor expansion $\sqrt{1-x} = 1 + \frac{x}{2} + \mathcal{O}(x^2)$ around $x_0 = 0$ results in

$$|\vec{x} - \vec{y}| \simeq |\vec{x}| - \frac{\vec{x}\cdot\vec{y}}{|\vec{x}|} = r_\sigma(p)^2 + \vec{k}'\cdot(q - \sigma) \ ,$$

where we defined $\vec{k}' = k\frac{\vec{x}}{|\vec{x}|}$, that points from the scattering center radially towards the detector and has the same norm as the wave vector $\vec{k}$. It follows:

$$-\frac{2m}{4\pi\hbar^2} \int\limits_{E_3} \frac{e^{ikr_p}}{r_p} V\psi \ dV \simeq -\frac{m}{2\pi\hbar^2} \int\limits_{E_3} \frac{e^{ikr_\sigma(p) - i\vec{k}'\cdot(q-\sigma)}}{r_\sigma(p) - \vec{k}'\cdot(q-\sigma)} V(q)\psi(q) \ dV_q \ ,$$

where we have made the q dependence explicit, to make the integration more transparent. In the case $r_\sigma(p) \to \infty$, the following approximation is asymptotically valid:

$$-\frac{2m}{4\pi\hbar^2} \int\limits_{E_3} \frac{e^{ikr_p}}{r_p} V\psi \ dV \simeq \frac{e^{ikr_\sigma(p)}}{r_\sigma(p)} f(\vec{k}', k, \psi)$$

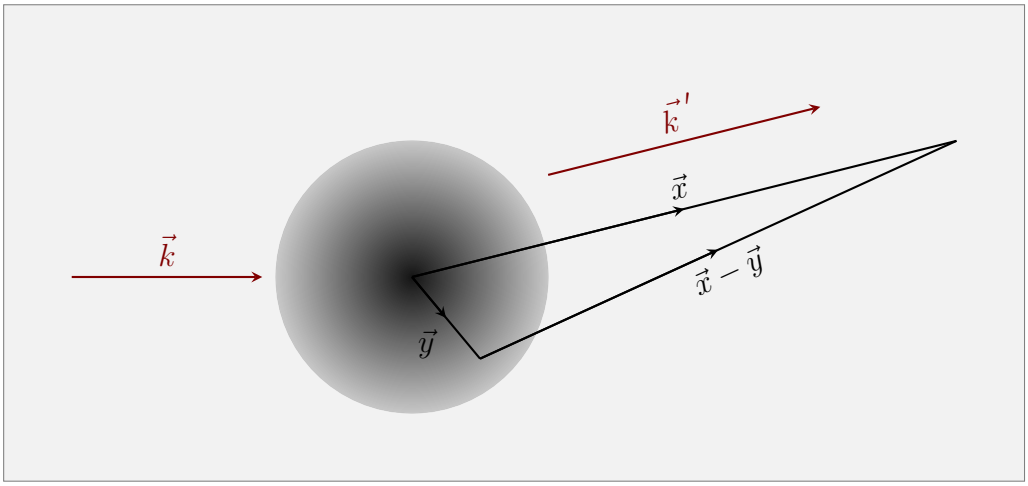


Figure 1.1.: Scattering setup

$$\text{with} \qquad f(\vec{k}', k, \psi) = -\frac{m}{2\pi\hbar^2} \int_{E_3} e^{-i\vec{k}'\cdot(q-\sigma)} V(q)\psi(q)\ dV_q \ .$$

Choosing σ as coordinate origin, the scattering amplitude becomes:

$$\boxed{f(\vec{k}', k, \psi) = -\frac{m}{2\pi\hbar^2} \int_{\mathbb{R}^3} e^{-i\vec{k}'\cdot\vec{y}} V(\vec{y})\psi(\vec{y})\ dy^3 = -\frac{m}{2\pi\hbar^2} \langle \vec{k}' | \widehat{V} | \psi \rangle \ .}$$

Theorem 1.3.1.
The differential cross section of plane waves is related to the scattering amplitude by

$$\frac{d\sigma}{d\Omega} \simeq |f(\vec{k}, \vec{k}', \psi)|^2 \ .$$

Proof 1.3.2.
To full wave function ψ is a super position of the incoming plane wave $\psi_i(\vec{x}) = e^{ikz}$ and the scattered wave $\psi_s = \psi - \psi_i$. We know that plane waves are solutions of the free Hamiltonian $\widehat{H}_0$, that is $\psi_0 = \psi_i$. We thus have:

$$\psi = e^{ikz} + \psi_s \simeq e^{ikz} + \frac{e^{ikr}}{r} f(\vec{k}, \vec{k}', \psi) \ .$$

With the definition of the probability current density 2-form $j = \frac{\hbar}{m} \star \mathfrak{Im}(\overline{\psi} \cdot d\psi)$ we get:

$$j_i = \frac{k\hbar}{m}[dx \wedge dy, R]$$

$$\text{and} \quad j_s = \frac{k\hbar}{m} \overline{f(\vec{k}, \vec{k}', \psi)} \cdot f(\vec{k}, \vec{k}', \psi)[\sin(\theta) d\theta \wedge d\phi, R] \ .$$

Using the definition of the differential cross section (definition 1.1.1), we get:

$$\frac{d\sigma}{d\Omega} = \overline{f(\vec{k}, \vec{k}', \psi)} \cdot f(\vec{k}, \vec{k}', \psi) = |f(\vec{k}, \vec{k}', \psi)|^2 \ .$$

□

1.4. Born approximation

To motivate the Born approximation we introduce a generalization of the geometric series for bounded operators:

Proposition 1.4.1.
Let $(X, ||\cdot||)$ be a Banach space and $T\colon X \to X$ a bounded operator. If the ***Neumann series*** $\sum_{k=0}^{\infty} T^k$ *converges with respect to the operator norm $||\cdot||_{op}$ then $(\mathbb{1} - T)$ is invertible and:*

$$(\mathbb{1} - T)^{-1} = \sum_{k=0}^{\infty} T^k \ .$$

Rewriting the Lippmann-Schwinger equation, we can obtain a form to use the Neumann series:

$$|\psi\rangle = (\mathbb{1} - \widehat{G}_0\widehat{V})^{-1}|\psi_0\rangle \ .$$

If $\widehat{G}_0\widehat{V}$ is bounded, $|\psi\rangle$ can be calculated from a series of operators, acting on $|\psi_0\rangle$:

$$|\psi\rangle = \sum_{n=0}^{\infty} \left(\widehat{G}_0\widehat{V}\right)^n \ |\psi_0\rangle \ .$$

In the special case $||\widehat{G}_0\widehat{V}||_{\text{op}} \ll 1$, the series can be terminated after the first two summands, called **first order Born approximation**:

$$\boxed{|\psi\rangle = |\psi_0\rangle + \widehat{G}_0\widehat{V}|\psi_0\rangle \ .}$$

If we assume the incoming wave to be a plane wave,

$$\langle\vec{y}|\psi_0\rangle = e^{i\vec{k}\cdot\vec{y}} \ ,$$

the scattering amplitude becomes essentially the Fourier transformation of the potential:

$$\boxed{f(\vec{k},\vec{k}') = -\frac{m}{2\pi\hbar^2}\int_{\mathbb{R}^3} e^{i(\vec{k}-\vec{k}')\cdot\vec{y}}V(\vec{y})\ dy^3 = -\frac{m}{2\pi\hbar^2}\langle\vec{k}-\vec{k}'|V\rangle \ .}$$

Lemma 1.4.2.
Let V be a spherical symmetrical potential $V(\vec{x}) = V(r)$. For elastic scattering, i.e. $|\vec{k}| = |\vec{k}'|$, the scattering amplitude can be calculated with

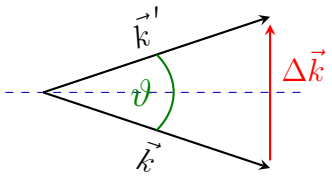
$$f(\vec{k},\vec{k}') = f(k,\vartheta) = -\frac{2m}{\hbar^2\Delta k}\int_0^{\infty} r\sin(\Delta k r)V(r)\ dr \ ,$$

where $\Delta k = 2k\sin(\vartheta/2)$.

Proof 1.4.3.
We define $\Delta\vec{k} = \vec{k}' - \vec{k}$. Choose coordinates such that $\Delta\vec{k}$ is directed alongside the z-axis. Calculating $f(\vec{k}, \vec{k}')$ in spherical coordinates results in:

$$\begin{aligned} f(\vec{k}, \vec{k}') &= -\frac{m}{2\pi\hbar^2} \int_{\mathbb{R}^3} e^{-i\Delta k r \cos(\theta)} V(r) r^2 \sin(\theta) dr \; d\phi \; d\theta \\ &= -\frac{2m}{\hbar^2} \int_0^\infty \int_{-1}^1 e^{-i\Delta k r \cos(\theta)} V(r) r^2 d(\cos(\theta)) \; dr \\ &= -\frac{2m}{\hbar^2 \Delta k} \int_0^\infty r \sin(\Delta k r) V(r) \; dr \; . \end{aligned}$$

The relation between Δk and k can be seen (trigonometry) from the following figure:



□

1.5. Scattering operator

Technically we have not considered scattering processes so far. Formally speaking, we have considered perturbations of the free Hamiltonian by a potential. That is, we have considered a stationary solution of the free Hamiltonian (plane wave) to find a solution of the full problem. A basic assumption to do so, is the time independence of the potential (or at least a slowly variation).

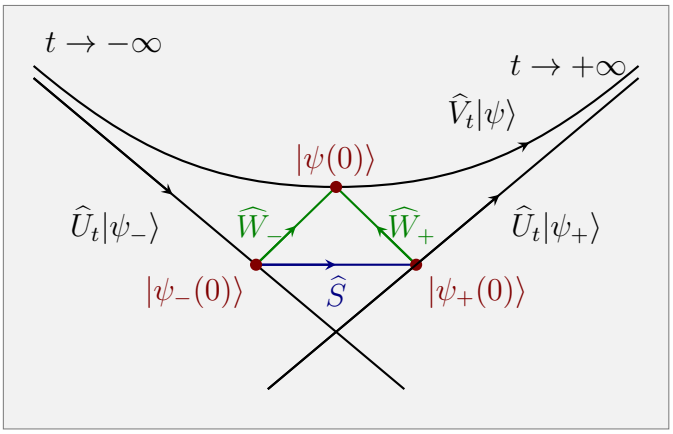


Figure 1.2.: Visualization of time dependent scattering. See [Zir10, p. 17] for the original

With scattering, in an intuitive meaning, we would associate the collision of two objects, or waves and fields for that matter, that change over time. The time dependence of the states can be expressed by a time evolution operator, such that

$$|\psi_0(t)\rangle = \widehat{U}_t|\psi_0\rangle \qquad \text{and} \qquad |\psi(t)\rangle = \widehat{V}_t|\psi\rangle \ .$$

From the Schrödinger equation we can see the equation for the full interaction time evolution operator $\widehat{V}_t$ and the interaction free time evolution operator $\widehat{U}_t$:

$$i\hbar\partial_t\widehat{V}_t = \widehat{H}\widehat{V}_t \qquad \text{and} \qquad i\hbar\partial_t\widehat{U}_t = \widehat{H}\widehat{U}_t \ .$$

Here the Hamiltonians need no longer be time-independent. The whole idea of the time-dependent scattering can be understood from figure 1.2:

Let $|\psi(t)\rangle$ be the state that describes the particle that is scattered. In the distant past, the particle will have been far away from the scattering potential and thus be moving freely essentially. Formally, we assume that there is an eigen state $|\psi_-(t)\rangle$ of $\widehat{H}_0$, such that

$$\lim_{t\to-\infty}|\psi_-(t)\rangle = \lim_{t\to-\infty}|\psi(t)\rangle \ .$$

Before we investigate this assumption any further, we will consider the distant future, as it imposes some difficulty. The state of the scattered particle may contain some component, that describes a bounded state (non zero projection on the space of bounded states). This component will not describe free motion even in the distant future. Since the norm is a continuous operator and thus limit and norm commute, our problem translates to

$$\begin{aligned}\lim_{t\to-\infty}\|\widehat{V}_t|\psi\rangle - \widehat{U}_t|\psi_-\rangle\| &= \lim_{t\to-\infty}\||\psi(t)\rangle - |\psi_-(t)\rangle\| \\ &= \|\lim_{t\to-\infty}|\psi(t)\rangle - \lim_{t\to-\infty}|\psi_-(t)\rangle\| = 0 \\ \text{but } \lim_{t\to+\infty}\|\widehat{V}_t|\psi\rangle - \widehat{U}_t|\psi_+\rangle\| &= \lim_{t\to+\infty}\||\psi(t)\rangle - |\psi_+(t)\rangle\| \\ &= \|\lim_{t\to+\infty}|\psi(t)\rangle - \lim_{t\to+\infty}|\psi_+(t)\rangle\| \neq 0 \ .\end{aligned}$$

To solve that problem, we will have to choose a different $|\psi\rangle_+$. To find the condition that determines $|\psi_+\rangle$, we pass to the non-limit case first.

Any time evolution operator is unitary:

Proof 1.5.1.
The adjoint operator satisfies $-i\hbar\partial_t\widehat{V}_t^\dagger = \widehat{V}_t^\dagger\widehat{H}$, since the Hamiltonian is self adjoint. Plugging this in, yields:

$$\begin{aligned}\partial_t\langle\widehat{V}_t\phi|\widehat{V}_t\psi\rangle = \partial_t\langle\phi|\widehat{V}_t^\dagger\widehat{V}_t|\psi\rangle &= \langle\phi|(\partial_t\widehat{V}_t^\dagger)\widehat{V}_t + \widehat{V}_t^\dagger\partial_t\widehat{V}|\psi\rangle \\ &= \langle\phi|\tfrac{i}{\hbar}\widehat{V}_t^\dagger\widehat{H}\widehat{V}_t - \tfrac{i}{\hbar}\widehat{V}_t^\dagger\widehat{H}\widehat{V}_t|\psi\rangle = 0 \ .\end{aligned}$$

But that means:

$$\begin{aligned}\langle\widehat{V}_t\phi|\widehat{V}_t\psi\rangle = \text{const.}(t) \quad \text{and} \quad \langle\widehat{V}_0\phi|\widehat{V}_0\psi\rangle = \langle\phi|\psi\rangle \\ \Rightarrow \qquad \langle\widehat{V}_t\phi|\widehat{V}_t\psi\rangle = \langle\phi|\psi\rangle \ .\end{aligned}$$

□

The unitarity has the following important properties for our purpose here:

$$\widehat{V}_t^\dagger \widehat{V}_t = \mathbb{1} \qquad \text{and} \qquad \|\widehat{V}_t|\psi\rangle\| = \||\psi\rangle\| = \|\widehat{V}_t^\dagger|\psi\rangle\| \ .$$

These properties allow to write

$$\|\widehat{V}_t|\psi\rangle - \widehat{U}_t|\psi_\pm\rangle\| = \||\psi\rangle - \widehat{V}_t^\dagger \widehat{U}_t|\psi_\pm\rangle\| \ .$$

Definition 1.5.2.
The **Møller operators** $W_\pm$ are defined by

$$W_\pm \lim_{t\to\pm\infty} \widehat{V}_t^\dagger \widehat{U}_t \ ,$$

if the limits exist.

The existence of the Møller operators is not guaranteed. From the physically reasoning above, we assume W_- to exist. The existence of W_+ however will be assumed as problem, that has to be checked for every scattering scenario independently. In the case of existence, the conditions for $|\psi_\pm\rangle$ are:

$$\||\psi\rangle - W_\pm|\psi_\pm\rangle\| = 0 \qquad \Leftrightarrow \qquad |\psi\rangle = W_\pm|\psi_\pm\rangle \ .$$

As can be seen in figure 1.2, the Møller operators map $|\psi_\pm\rangle$ to the scattered state $|\psi\rangle$ at the moment of scattering. This motivates the following definition:

Definition 1.5.3.
If range$(\widehat{W}_-) \subset$ domain$(\widehat{W}\dagger_+)$, the **scattering operator** $\widehat{S}$ is defined by

$$\widehat{S} = \widehat{W}_+^\dagger \widehat{W}_- \ .$$

The scattering operator maps $|\psi_-\rangle$, that describes the state of scattered particle in the distant part, to $|\psi_+\rangle$, that describes the free components of the state of the particle in the distant future.

2
Second quantization

The name second quantization, motivated by its historic origin, can be misleading. A more appropriate name for this chapter would be "many particle systems of fermions and bosons". Before the formulation of second quantization is introduced, we focus on distinguishable particles first. In doing so, the Copenhagen postulate becomes a consequence of the quantum mechanical description of measurement processes. This chapter is based on [Zir10] for the mathematical background and [Sch08] for the connection to the textbook conventions.

2.1. Distinguishable particles

It is easiest to consider non-interacting particles first. Since a fundamental principle of quantum mechanics is the indistinguishability of identical particles, we consider different particles here. Most of this section can be found in [Sch13, Kapitel 20] in a different mathematical formulation. However, the part about the measurement process follows [RW08] closely.

2.1.1. Tensor Hilbert spaces and Dirac notation

Consider two systems, described by the Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively. To describe both systems together, the tensor space $\mathcal{H}_1 \otimes \mathcal{H}_2$ is used. With that information, from a mathematical point of view (see chapter C), all has been said about the Hilbert space of non interacting distinguishable particles. However, the contemporary literature mostly uses the Dirac notation, which needs further explanation.

Let $|\phi\rangle_1 \in \mathcal{H}_1$ and $|\psi\rangle_2 \in \mathcal{H}$ be two states of their respective systems. These states define a state in $\mathcal{H}_1 \otimes \mathcal{H}_2$:

$$|\phi\rangle_1 |\psi\rangle_2 := |\phi\rangle_1 \otimes |\psi\rangle_2 \in \mathcal{H}_1 \otimes \mathcal{H}_2 \ .$$

Furthermore it is common to not only drop the tensor product, but also the indices, writing $|\phi\rangle|\psi\rangle$, as long as no ambiguities arise. The bra-states ${}_1\langle\phi|\,{}_2\langle\psi|$ are defined, using the isomorphism

$$(\mathcal{H}_1 \otimes \mathcal{H}_2)^* \simeq \mathcal{H}_1^* \otimes \mathcal{H}_2^* \ .$$

It would also be possible to use $(\mathcal{H}_1 \otimes \mathcal{H}_2)^* \simeq \mathcal{H}_2^* \otimes \mathcal{H}_1^*$. However caution is necessary, when using both, not to intermingle isomorphy and equality.

The hermitian scalar product of the tensor space defined by $(|\phi'\rangle \otimes |\psi'\rangle, |\phi\rangle \otimes |\psi\rangle) = \langle\phi'|\phi\rangle \cdot \langle\psi'|\psi\rangle$ translates to

$${}_1\langle\phi'|\,{}_2\langle\psi'||\phi\rangle_1|\psi\rangle_2 = \langle\phi'|\phi\rangle \cdot \langle\psi'|\psi\rangle$$

in the Dirac notation. *Loosely speaking, bra-vectors and ket-vectors only connect, if they belong to the same Hilbert space, otherwise they can pass through each other.* This rule also holds true for the interpretation of tensors as multilinear maps:

$$(|\phi\rangle \otimes |\psi\rangle)(\langle\psi'|) = \langle\psi'|\psi\rangle|\phi\rangle = {}_2\langle\psi'||\phi\rangle_1|\psi\rangle_2$$

$$\text{and} \quad (\langle\phi'| \otimes \langle\psi'|)(|\phi\rangle) = \langle\psi'|\langle\phi'|\phi\rangle = {}_1\langle\phi'|\,{}_2\langle\psi'||\phi\rangle_1 \ .$$

If the systems do not interact, the probability for the first system to be in the state $|\chi\rangle$ and the second system to be in the state $|\xi\rangle$, should be the product of the individual properties. Indeed , we find:

$$\begin{aligned} P(\chi,\xi) &= |\langle\chi| \otimes \langle\xi||\phi\rangle \otimes |\psi\rangle|^2 = |\langle\chi|\phi\rangle \cdot \langle\xi|\psi\rangle|^2 = |\langle\chi|\phi\rangle|^2 \cdot |\langle\xi|\psi\rangle|^2 \\ &= P_1(\chi) \cdot P_2(\xi) \ . \end{aligned}$$

Lemma C.1.10 allows to construct operators for $\mathcal{H} \coloneqq \mathcal{H}_1 \otimes \mathcal{H}_2$ from operators $A_1 \in \mathrm{End}(\mathcal{H}_1)$ and $A_2 \in \mathrm{End}(\mathcal{H}_2)$:

$$\mathbf{A}_1 = A_1 \otimes \mathbb{1} \in \mathrm{End}(\mathcal{H}) \qquad \text{and} \qquad \mathbf{A}_2 = \mathbb{1} \otimes A_2 \in \mathrm{End}(\mathcal{H}) \ .$$

To check, that $\mathbb{A}_1$ is the correct transcription of A_1, we can calculate the expectation value:

$$\begin{aligned} \langle\mathcal{A}_1\rangle_\Psi = \langle\Psi|\mathcal{A}_1|\Psi\rangle &= \langle\Psi_1|\langle\Psi_2|A_1 \otimes \mathbb{1}|\Psi_1\rangle|\Psi_2\rangle = \langle\Psi_1|A_1|\Psi_1\rangle \cdot \underbrace{\langle\Psi_2|\mathbb{1}|\Psi_2\rangle}_{=1} \\ &= \langle A_1\rangle_{\Psi_1} \ . \end{aligned}$$

2.1.2. Density matrix

Superposition of states fails to represent classical probabilities. Consider for example a process, that creates an equal amount of spin up $|\uparrow\rangle$ and spin down $|\downarrow\rangle$ electrons. Encountering one electron, the probability is 0.5 for each state. However, the superposition $|\uparrow\rangle + |\downarrow\rangle$ describe an electron with spin in x-direction (up to normalization). Hence the need for a new quantity, to describe classical probabilities.

Definition 2.1.1.
Let Ψ_j be possible normalized states with probabilities p_j. The **density operator** ρ is defined by

$$\rho = \sum_j p_j|\Psi_j\rangle\langle\Psi_j| \coloneqq \sum_j p_j|\Psi_j\rangle \otimes \langle\Psi_j| \in \mathcal{H} \otimes \mathcal{H}^* \ .$$

The expectation value of an Operator $A \in \mathrm{End}(\mathcal{H})$ can be calculated using the trace:

$$\begin{aligned} \langle A\rangle &= \sum_j p_j\langle\Psi_j|A|\Psi_j\rangle = \sum_{n,j} p_j\langle\Psi_j|n\rangle\langle n|A|\Psi_j\rangle \\ &= \sum_n \langle n|\left(A\sum_j p_j|\Psi_j\rangle\langle\Psi_j|\right)|n\rangle = \sum_n \langle n|A\rho|n\rangle = tr(A\rho) \end{aligned}$$

$$\Rightarrow \qquad \boxed{\langle A\rangle = tr(A\rho)\ .}$$

A property of density operators is to have unit trace $\mathrm{tr}(\rho) = 1$. This can be seen from the previous calculation using $A = \mathbb{1}$. It could seem annoying, that in calculating the trace, a Hilbert basis was chosen. However, the next lemma shows, that the trace does not depend on such a choice.

Lemma 2.1.2.
Let $A \in End(\mathcal{H})$ be an operator defined by

$$A = \sum_k p_k |\psi_k\rangle\langle\phi_k| = \sum_k p_k |\psi_k\rangle \otimes \langle\phi_k|\ ,$$

then the trace $tr(A)$ can be calculated as follows (and thus does not depend upon the choice of a Hilbert basis):

$$tr(A) = \sum_k p_k \langle\phi_k|\psi_k\rangle\ .$$

Proof 2.1.3.
Let $\{\langle j|\}_j$ be an arbitrary Hilbert basis. Using the completeness relation we see, that:

$$\begin{aligned} tr(A) &= \sum_j A_{jj} = \sum_j \langle j|A|j\rangle = \sum_{j,k} p_k \langle j|\psi_k\rangle\langle\phi_k|j\rangle \\ &= \sum_k p_k \langle\phi_k| \left(\sum_j |j\rangle\langle j|\right) |\psi_k\rangle = \sum_k p_k \langle\phi_k|\psi_k\rangle \end{aligned}$$

□

Theorem 2.1.4.
*The time evolution of $\rho(t)$ is determined by the **von Neumann equation**:*

$$i\hbar\partial_t \rho(t) = -[\rho(t), H]\ .$$

Proof 2.1.5.
The Schrödinger equation yields the following equations:

$$i\hbar\partial_t |\psi_n\rangle = H|\psi_n\rangle \qquad \Rightarrow \qquad -i\hbar\partial\langle\psi_n| = \langle\psi_n|H\ .$$

Combining these equations results in the claim:

$$i\hbar\partial_t \rho = \sum_n p_n (H|\psi_n\rangle\langle\psi_n| - |\psi_n\rangle\langle\psi_n|H) = [H, \rho]\ .$$

□

For time independent Hamilton operators, a formal solution is given by $\rho(t) = e^{-iHt/\hbar}\rho(0)\ e^{iHt/\hbar}$.

Lemma 2.1.6.
The density operator ρ is hermitian and positive semi definite.

Proof 2.1.7.

i) $\langle\phi|\widehat{\rho}|\chi\rangle = \sum_n p_n\langle\phi|\psi_n\rangle\langle\psi_n|\chi\rangle = \overline{\sum_n p_n\langle\chi|\psi_n\rangle\langle\psi_n|\phi\rangle} = \overline{\langle\chi|\widehat{\rho}|\phi\rangle} = \langle\widehat{\rho}\phi|\chi\rangle$,

ii) $\langle\phi|\widehat{\rho}|\phi\rangle = \sum_n p_n|\langle\phi|\psi_n\rangle|^2 \geq 0 \qquad \forall\ |\phi\rangle \in \mathcal{H}$.

□

Definition 2.1.8.
A state described by a density operator is called **pure**, if there is a $|\Psi\rangle \in \mathcal{H}$, such that $\rho = |\Psi\rangle\langle\Psi|$, otherwise it is called **mixed** .

Lemma 2.1.9.
A state is pure, if and only if $tr(\rho^2) = 1$. Conversely, a state is mixed, if and only if $tr(\rho^2) < 1$.

Proof 2.1.10.

1. If ρ is pure, then:

$$\rho^2 = |\Psi\rangle\langle\Psi|\Psi\rangle\langle\Psi| = |\Psi\rangle\langle\Psi| = \rho \qquad \Rightarrow \qquad tr(\rho^2) = tr(\rho) = 1 \ .$$

2. Since ρ is positive semi definite, there is a Hilbert basis, such that $\rho = \sum_n p_n|\psi_n\rangle\langle\psi_n|$ with $p_n \geq 0$. From $tr(\rho) = 1$ it follows, that

$$\sum_n p_n = 1 \qquad \Rightarrow \qquad tr(\rho^2) = \sum_n p_n^2 \leq \sum_n p_n = 1 \ .$$

Assuming $tr(\rho^2) = 1$, then:

$$0 = \sum_n p_n - \sum_n p_n^2 = \sum_n p_n(1 - p_n) \qquad \Rightarrow \qquad p_n = 0 \text{ or } p_n = 1 \ .$$

Yet, since $tr(\rho) = 1$, there has to be an n_0 such that $p_{n_0} = 1$, and thus:

$$p_n = \begin{cases} 1 & n = n_0 \\ 0 & n \neq n_0 \end{cases}$$

This means $\rho = |\psi_{n_0}\rangle\langle\psi_{n_0}|$, which would mean, that the state is pure. □

So far, no further structure was imposed on the Hilbert space $\mathcal{H}$. For that reason, the density operator carries over seamlessly to many particle systems. For example, take

the Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ from the previous subsection. With general states $|\Psi_j\rangle = \sum_{m,n} |\psi_{jm}\rangle|\phi_{jn}\rangle$, the density operator becomes

$$\rho = \sum_{j,m,m',n,n'} p_j |\psi_{jm}\rangle|\phi_{jn}\rangle\langle\psi_{jm'}|\langle\phi_{jn'}| \in \mathcal{H}_1^* \otimes \mathcal{H}_2^* \otimes \mathcal{H}_1 \otimes \mathcal{H}_2 \ ,$$

where all sums are finite. To define a partial trace operator

$$T_1 \colon \mathcal{H}_1^* \otimes \mathcal{H}_2^* \otimes \mathcal{H}_1 \otimes \mathcal{H}_2 \to \mathcal{H}_2^* \otimes \mathcal{H}_2 \ ,$$

that acts as trace on the Hilbert space $\mathcal{H}_1$, we interpret bra-vectors as functionals on the Hilbert space and ket-vectors as functionals on the dual space. Choosing an arbitrary Hilbert basis $\{|\ell\rangle\}$ on $\mathcal{H}_1$, we define

$$T_1 = \sum_\ell \langle\ell| \otimes \mathbb{1} \otimes |\ell\rangle \otimes \mathbb{1} \ .$$

The action of T_1 on ρ yields:[1]

$$\begin{aligned} T_1\rho &= \sum_{j,m,n} p_j \sum_\ell \langle\psi_{jm'}|\ell\rangle\langle\ell|\psi_{jm}\rangle|\phi_{jn}\rangle\langle\phi_{jn'}| \\ &= \sum_j p_j \left(\sum_{m,m'} \langle\psi_{jm'}|\psi_{jm}\rangle\right) \left(\sum_{n,n'} |\phi_{jn}\rangle\langle\phi_{jn'}|\right) \ . \end{aligned}$$

The normalization condition for $|\Psi_j\rangle$ results in

$$1 = \langle\Psi_j|\Psi_j\rangle = \left(\sum_{m,m'} \langle\psi_{jm'}|\psi_{jm}\rangle\right) \left(\sum_{n,n'} \langle\phi_{jn}|\phi_{jn'}\rangle\right) \ .$$

This allows to choose coefficient such that $\sum_{m,m'}\langle\psi_{jm'}|\psi_{jm}\rangle = 1$. Also, by the multilinearity of the tensor product:

$$\sum_{n,n'} |\phi_{jn}\rangle\langle\phi_{jn'}| = \left(\sum_n |\phi_{jn}\rangle\right) \otimes \left(\sum_{n'} \langle\phi_{jn'}|\right) \coloneqq |\Phi_j\rangle\langle\Phi_j| \ ,$$

with $|\Phi_j\rangle \in \mathcal{H}_2$. Finally we see, that

$$T_1\rho = \sum_j p_j |\Phi_j\rangle\langle\Phi_j|$$

is a density operator on $\mathcal{H}_2$.

Definition 2.1.11.
Let ρ be a density operator on $\mathcal{H}_1 \otimes \mathcal{H}_2$, then $T_1\rho$ is called **reduced density operator** on $\mathcal{H}_2$. In the same way, the reduced density operator $T_2\rho$ is defined on $\mathcal{H}_1$.

[1] To exchange the infinite sum over ℓ with the other sums, we had to use, that the other sums are finite.

2.1.3. Measurement process

A measurement can be understood as two quantum systems, the one to be observed $\mathcal{H}_O$, and the measurement system $\mathcal{H}_S$, i.e. the measurement device. The full system is as explained above described by $\mathcal{H} = \mathcal{H}_O \otimes \mathcal{H}_M$. The full Hamiltonian, that determines the time evolution is assumed to be as sum of the observation system Hamiltonian H_O, the measurement system H_M and an interaction Hamiltonian H_I

$$H = H_O + H_M + H_I \ .$$

The interaction Hamiltonian, as the same suggests, describes the interaction between the observation system and the measurement system, allowing measurements to be taken. An ideal measurement can be characterized as instantaneous, such that no time development has to be taken into account. Also, a measurement is always defined w.r.t a physical quantity, that means an Operator A. Assuming $|\alpha_n\rangle$ to be a Hilbert eigen basis of A, the interaction Hamiltonian should look like

$$H_I = \sum_n |\alpha_n\rangle\langle\alpha_n| \otimes H_M(\alpha_n) \ .$$

The measurement process (time development) of a state $|\Psi\rangle = |\alpha_n\rangle|\chi\rangle$ creates a state $|\Psi'\rangle = |\alpha_n\rangle|\chi(\alpha_n)\rangle$, where $|\chi(\alpha_n)\rangle$ is a characteristic state in the measurement system, corresponding to the state $|\alpha_n\rangle$. To idealize the measurement further, the characteristic states are supposed to have no overlap with each other, such that they can be distinguished perfectly even long after the measurement process:

$$\langle\chi(\alpha_m)|\chi(\alpha_n)\rangle = \delta_{mn} \ .$$

Just before a measurement, the density operator can be written as follows:

$$\begin{aligned}\rho = |\Psi\rangle\langle\Psi| &=: |\psi\rangle|\chi\rangle\langle\psi|\langle\chi| = \sum_{n,m} |\psi\rangle|\alpha_n\rangle\langle\alpha_n|\chi\rangle\langle\psi|\langle\chi|\alpha_m\rangle\langle\alpha_m| \\ &=: \sum_{n,m} c_{m,n}\ |\psi\rangle|\alpha_n\rangle\langle\psi|\langle\alpha_m| \ ,\end{aligned}$$

with $c_{m,n} = \langle\alpha_n|\chi\rangle \cdot \langle\chi|\alpha_m\rangle$. After the measurement, a similar calculation shows, that the density operator is

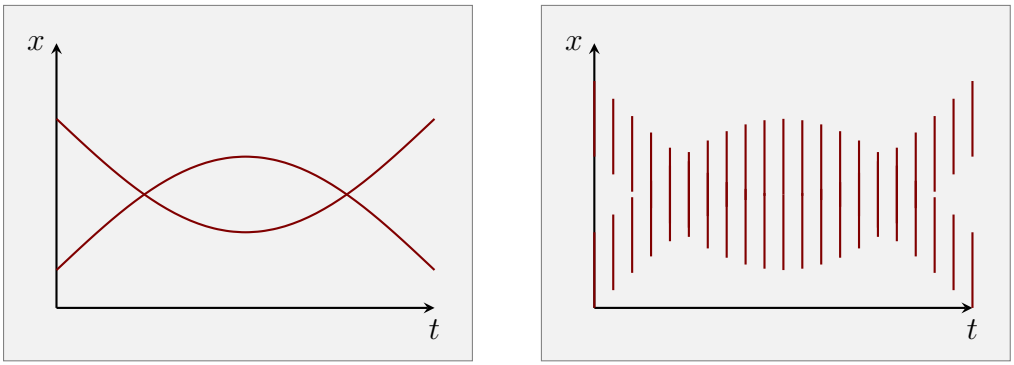
$$\rho' = \sum_{n,m} c_{m,n}\ |\alpha_n\rangle|\chi(\alpha_n)\rangle\langle\alpha_m|\langle\chi(\alpha_m)| \ .$$

To see the effects of the measurement on the observation system, the reduced density operator on $\mathcal{H}_O$ is calculated:

$$\begin{aligned}T_2\rho &= \sum_{j,m,n} c_{m,n}\ \langle j|\chi(\alpha_n)\rangle\langle\chi(\alpha_m)|j\rangle|\alpha_n\rangle\langle\alpha_m| \\ &\overset{\sum_j |j\rangle\langle j| \mathbb{1}}{=} \sum_{m,n} c_{m,n}\ \langle\chi(\alpha_m)|\chi(\alpha_n)\rangle|\alpha_n\rangle\langle\alpha_m| \\ &\underset{\langle\psi,\alpha_m|\psi,\alpha_n\rangle=\delta_{m,n}}{=} \sum_m c_{m,m}\ |\alpha_m\rangle\langle\alpha_m|\end{aligned}$$

After the measurement, the off-diagonal terms in the observation system have vanished. The state has become an ensemble of possible eigenstates with probabilities $c_{m,m} = |\langle\alpha_m|\chi\rangle|^2$. The meaning of density operators as realizations of statistical ensembles suggests the Copenhagen postulate, as it does not describe superpositions, but rather classical probabilities. This means, that after the measurement, the state in the observation system, has collapsed in one definite eigenstate. Yet which one, is a matter of the probability $c_{m,m} = |\langle\alpha_m|\chi\rangle|^2$.

2.2. Identical particles



(a) (b)

Figure 2.1.: In the first case, the classical picture is illustrated. Each trajectory has infinite precise spacial resolution, such that particles can be distinguished easily. In the second case, the quantum mechanical case with high position uncertainty is depicted. Here it is impossible to distinguish the particles, after they have come into reach of each other, i.e. forming a system.

As mentioned before, identical particles cannot be distinguished. Put differently, it is not possible to label particles of the same sort in any way, that allows to identify them later on. The first problem of such an undertaking is, that a particle type is described by a set of quantum properties, that is the same for each individual particle. Each electron has the same charge, the same mass and the same total spin. This is a behavior, not common in classical mechanics, where there can always be found properties that vary even between identical objects. One could assume, that constantly observing a system would allow to keep track and distinguish each particle. However, the uncertainty relations of several observables make this impossible. A high spacial resolution leads to a high uncertainty of the momentum, i.e. the direction and magnitude of movement. After the moment, one assigned each particle a number (not physically, which is impossible, but mentally), it is not clear, which has moved in which direction. The assignment is of no use for that matter.

On the other hand, having a high momentum resolution results in a high position uncertainty. It is no longer possible to discern the particles separately, such that any assignment of labels is of no use again. This is illustrated in figure 2.1.

2.2.1. Symmetry postulate

A consequence of the indistinguishability of identical particles is, that any exchange of these particles must not have any change of physical properties. Any description of a system of particles must satisfy this rule.

We consider a system of n identical particles. In the last section we have seen, that the Hilbert space for distinguishable particles is a tensor product $\mathcal{H}^{\otimes n}$. Hence it is reasonable, that the Hilbert space for identical particles is a subspace of $\mathcal{H}^{\otimes n}$. To find this subspace, we begin by considering general states

$$|\psi_1\rangle \dots |\psi_n\rangle \coloneqq |\psi_1\rangle \otimes \dots \otimes |\psi_n\rangle \in \mathcal{H}^{\otimes n} \ .$$

A permutation operator P can be defined by

$$P(|\psi_1\rangle \dots |\psi_n\rangle) \coloneqq |\psi_{P(1)}\rangle \dots |\psi_{P(n)}\rangle \ .$$

Before we continue to find restrictions for the states of identical particles, we show some properties of the permutation operator:

Corollary 2.2.1.
Any transposition operator P_{ij} is an involution, i.e. $P_{ij}^2 = \mathbb{1}$. Any permutation operator is unitary and all transposition operators are also hermitian.

Proof 2.2.2.
The first claim is obvious from permutations. The second claim follows from the definition of the hermitian scalar product $(\cdot,\cdot)$ on tensor spaces:

$$\begin{aligned}(P(|\phi_1\rangle \dots |\phi_n\rangle), P(|\psi_1\rangle \dots |\psi_n\rangle)) &= (|\phi_{P(1)}\rangle \dots |\phi_{P(n)}\rangle, |\psi_{P(1)}\rangle \dots |\psi_{P(n)}\rangle) \\ &= \prod_{j=1}^{n} \langle \phi_{P(j)} | \psi_{P(j)} \rangle = \prod_{j=1}^{n} \langle \phi_j | \psi_j \rangle \\ &= (|\phi_1\rangle \dots |\phi_n\rangle, |\psi_1\rangle \dots |\psi_n\rangle) \ .\end{aligned}$$

The last claim follows immediately from the previous two. □

Lemma 2.2.3.
The eigen values p of a permutation operator can only be $p = +1$ and $p = -1$. Also, if p is the eigen value of the permutation operator P for the eigen state $|\Psi\rangle$, then $p' = p$ for all other permutation operators P' acting on $|\Psi\rangle$.

Proof 2.2.4.
Let P_{ij} be a transposition operator, then

$$P_{ij}^2 |\Psi\rangle = \mathbb{1}|\Psi\rangle = (a_{ij})^2 |\Psi\rangle \quad \Rightarrow \quad a_{ij} = \pm 1 \ .$$

Let $P_{12}|\Psi\rangle = a_{12}|\Psi\rangle$, the $|\Psi\rangle$ is an eigen state to the eigen value $a_{ij} = a_{12}$ of all other transposition operators P_{ij} with $i \neq j$. This can be seen by writing P_{ij} as follows:

$$P_{ij} = P_{1j} \circ P_{2i} \circ P_{12} \circ P_{2i} \circ P_{1j} \quad \Rightarrow \quad a_{ij} = a_{1j}^2 \cdot a_{2i}^2 \cdot a_{12} = a_{12} \ .$$

Since all permutations can be written as product of transpositions, it follows that $P|\Psi\rangle = a_{12}|\Psi\rangle$ for all permutation operators. □

The tensor space $\mathcal{H}^{\otimes n}$ can be written as direct sum of a symmetric and an antisymmetric space $\mathcal{H}^{\otimes n} = S^n(\mathcal{H}) \oplus \bigwedge^n(\mathcal{H})$:

$$P_{ij}|\Psi\rangle = |\Psi\rangle \quad \forall\ |\Psi\rangle \in S^n(\mathcal{H}) \qquad \text{and} \qquad P_{ij}|\Psi\rangle = -|\Psi\rangle \quad \forall\ |\Psi\rangle \in \bigwedge\nolimits^n(\mathcal{H}) \ .$$

There could be other subspaces, that mix symmetric states and antisymmetric states. However, observations show, that this is not the case. Hence we postulate:

symmetry postulate

States of identical particles are either completely symmetric or completely antisymmetric. Particles with completely symmetric states are called **bosons** and particles with completely antisymmetric states are called **fermions**.

2.2.2. (Anti-)Symmetrization and scalar products

The symmetric and the wedge product can be constructed from tensor products. For that purpose we define the (anti-)symmetrization operators:

$$S_+ = \sum_{P \in \Sigma_n} P \qquad \text{and} \qquad S_- = \sum_{P \in \Sigma_n} \mathrm{sgn}(P) \cdot P \ .$$

Lemma 2.2.5.
The symmetrization operator S_+ symmetrizes and the antisymmetrization operator S_- antisymmetrizes states in $\mathcal{H}^{\otimes n}$.

Proof 2.2.6.
Consider a transposition operator P_{ij}, then:

$$\begin{aligned} P_{ij}(S_+|\Psi\rangle) = P_{ij} \sum_{P \in \Sigma_n} P|\Psi\rangle = \sum_{P \in \Sigma_n} (P_{ij} \circ P)|\Psi\rangle = \sum_{P \in \Sigma_n} P'(P)|\Psi\rangle \\ = \sum_{P' \in \Sigma_n} P'|\Psi\rangle = S_+|\Psi\rangle \ . \end{aligned}$$

In the same way:

$$\begin{aligned} P_{ij}(S_-|\Psi\rangle) = \sum_{P \in \Sigma_n} \mathrm{sgn}(P) \cdot (P_{ij} \circ P)|\Psi\rangle = \sum_{P \in \Sigma_n} -\mathrm{sgn}(P_{ij} \circ P) \cdot (P_{ij} \circ P)|\Psi\rangle \\ = -\sum_{P' \in \Sigma_n} \mathrm{sgn}(P') \cdot P'|\Psi\rangle = -S_-|\Psi\rangle \ . \end{aligned}$$

□

The operators $S_\pm$ allow to define a symmetric and a wedge product, with a normalization factor that is rather uncommon, but necessary to obtain the right scalar product:

$$|\psi_1 \vee \ldots \vee \psi_n\rangle \equiv |\psi_1\rangle \vee \ldots \vee |\psi_n\rangle := \frac{1}{\sqrt{n!}} S_+(|\psi_1\rangle \otimes \ldots \otimes |\psi\rangle_n) \in S^n(\mathcal{H}) \ ,$$
$$|\psi_1 \wedge \ldots \wedge \psi_n\rangle \equiv |\psi_1\rangle \wedge \ldots \wedge |\psi_n\rangle := \frac{1}{\sqrt{n!}} S_-(|\psi_1\rangle \otimes \ldots \otimes |\psi\rangle_n) \in \bigwedge^n(\mathcal{H}).$$

Lemma 2.2.7.
The scalar product induced by the scalar product on $\mathcal{H}^{\otimes n}$, has the following form on $S^n(\mathcal{H})$ and $\bigwedge^n(\mathcal{H})$:

$$\langle \psi_1 \vee \ldots \vee \psi_n | \phi_1 \vee \ldots \vee \phi_n \rangle = \sum_{P \in \Sigma_n} \prod_{j=1}^{n} \langle \psi_j | \phi_{P(j)} \rangle$$

and

$$\langle \psi_1 \wedge \ldots \wedge \psi_n | \phi_1 \wedge \ldots \wedge \phi_n \rangle = \det(\langle \psi_i | \phi_j \rangle)_{ij} \ .$$

Proof 2.2.8.
Using the definition of $\vee$ we find:

$$\langle \psi_1 \vee \ldots \vee \psi_n | \phi_1 \vee \ldots \vee \phi_n \rangle = \frac{1}{n!} \sum_{P,P' \in \Sigma_n} \prod_{j=1}^{n} \langle \psi_{P'(j)} | \phi_{P(j)} \rangle = \sum_{P \in \Sigma_n} \prod_{j=1}^{n} \langle \psi_j | \phi_{P(j)} \rangle \ .$$

The second equality used, that the sum over the permutations P' creates $n!$-times the same combinations, since the product is commutative.

For the scalar product of $\bigwedge^n(\mathcal{H})$ the Leibniz-formula of the determinant will be used:

$$\det(\langle \psi_i | \phi_j \rangle)_{ij} \ . = \sum_{P \in \Sigma_n} \operatorname{sgn}(P) \prod_{j=1}^{n} \langle \psi_j | \phi_{P(j)} \rangle \ .$$

$$\begin{aligned} \Rightarrow \quad \langle \psi_1 \wedge \ldots \wedge \psi_n | \phi_1 \wedge \ldots \wedge \phi_n \rangle &= \frac{1}{n!} \sum_{P,P' \in \Sigma_n} \operatorname{sgn}(P')\operatorname{sgn}(P) \prod_{j=1}^{n} \langle \psi_{P'(j)} | \phi_{P(j)} \rangle \\ &= \frac{1}{n!} \sum_{P' \in \Sigma_n} \operatorname{sgn}(P') \prod_{j=1}^{n} \det(\langle \psi_i | \phi_j \rangle)_{ij} = \det(\langle \psi_i | \phi_j \rangle)_{ij} \ . \end{aligned}$$

Because the determinant is alternating, the sign of $\operatorname{sgn}(P')$ is canceled out, such that $n!$-times the same term is summed over, resulting in the last equality. □

2.2.3. Representation on $L^2(\mathbb{R}^3)$

Independent, if we are looking for a representation in the position, or in the momentum space, in both cases we need a representation on $L^2(\mathbb{R}^3)$. All we need to do so, is a definition for the symmetric and the wedge product. To stay consistent with the

previous subsection, we define the coefficients different to the literature with a square root. For $\Psi_p \in S^p(L^2(\mathbb{R}^3))$ or $\Psi_p \in \bigwedge^p(L^2(\mathbb{R}^3))$ and analogously Ψ_q we define:

$$\Phi_p \vee \Psi_q(\vec{x}_1, ..., \vec{x}_{p+q}) \coloneqq \frac{1}{\sqrt{p! \cdot q!}} \sum_{P \in \Sigma_{p+q}} \Phi_p(\vec{x}_{P(1)}, ..., \vec{x}_{P(p)}) \cdot \Psi_q(\vec{x}_{P(p+1)}, ..., \vec{x}_{P(p+q)}) \ ,$$

$$\Phi_p \wedge \Psi_q(\vec{x}_1, ..., \vec{x}_{p+q}) \coloneqq \frac{1}{\sqrt{p! \cdot q!}} \sum_{P \in \Sigma_{p+q}} \operatorname{sgn}(P) \cdot \Phi_p(\vec{x}_{P(1)}, ..., \vec{x}_{P(p)}) \cdot \Psi_q(\vec{x}_{P(p+1)}, ..., \vec{x}_{P(p+q)}) \ .$$

2.3. Fock space

So far the description of identical particles lacks one principle of modern particle physics, the creation and annihilation of particles. To fix this, we pass to the so called **Fock spaces** for bosons $\mathcal{F}_+$ and fermions $\mathcal{F}_-$:

$$\mathcal{F}_+(\mathcal{H}) = \bigoplus_{N=0}^{\infty} S^N(\mathcal{H}) = S(\mathcal{H}) \qquad \mathcal{F}_-(\mathcal{H}) = \bigoplus_{N=0}^{\infty} \bigwedge\nolimits^N(\mathcal{H}) = \bigwedge(\mathcal{H}) \ .$$

The direct sum shows a grading of the symmetric and the exterior algebra. The hermitian scalar product of the last subsection carries over naturally because of this grading:

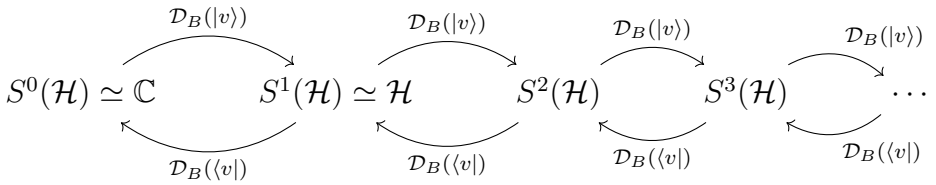
$$\langle c + v + v \vee w + u \vee v \vee w + ... | c' + v' + v' \vee w' + u' \vee v' \vee w' + ... \rangle$$

$$= \overline{c} \cdot c' + \langle v | v' \rangle_{\mathcal{H}} + \langle v \vee w | v' \vee w' \rangle_{S^2(\mathcal{H})} + ...$$

Although the descriptions of bosons and fermions share many properties, they are discussed in different subsections for a clearer exposition, at the cost of redundancy. The following subsections make use of the section D.4.

2.3.1. Bosons

The ground state, i.e. the state without any particle, is called **vacuum state** $|0\rangle \equiv 1 \in S^0(\mathcal{H}) \simeq \mathbb{C}$. Let $|v\rangle \in \mathcal{H}$ and $\langle \varphi | \in \mathcal{H}^*$. The bosonic Fock space representation allows to understand these elements as linear maps on $S(\mathcal{H})$ that increase and decrease the degree respectively. Physically speaking, the representations $\mathcal{D}_B(|v\rangle)$ and $\mathcal{D}_B(\langle v|)$ create a particle in the state $|v\rangle$ and annihilate a particle in the dual state $|v\rangle$ respectively:



Definition 2.3.1.
Let $\{|e_j\rangle\}$ be a Hilbert basis of $\mathcal{H}$ and $\{\langle e_j|\}$ a dual Hilbert basis. The **creation/ annihilation operators** $a_j^\dagger / a_j$ are defined by

$$a_j^\dagger \coloneqq \mathcal{D}_B(|e_j\rangle) \qquad \text{and} \qquad a_j \coloneqq \mathcal{D}_B(\bar{e_j}) \ .$$

Corollary 2.3.2.
The bosonic creation/annihilation operators satisfy the ***canonical commutator relations****:*

$$[a_j^\dagger, a_k^\dagger] = 0 \qquad [a_k, a_k] = 0 \qquad [a_k, a_j^\dagger] = \delta_{jk}\mathbb{1} \qquad \forall\ j,k\ .$$

Proof 2.3.3.
This is just a special case of theorem D.4.6. □

Definition 2.3.4.
The **occupation number states** are defined as follows:

$$|n_1, n_n, \ldots, n_\ell, \ldots\rangle \coloneqq \frac{(a_1^\dagger)^{n_1}}{\sqrt{n_1!}} \frac{(a_2^\dagger)^{n_2}}{\sqrt{n_2!}} \ldots \frac{(a_\ell^\dagger)^{n_\ell}}{\sqrt{n_\ell!}} \ldots |0\rangle\ .$$

Using the convention $|v\rangle^n = |v\rangle \vee \ldots \vee |v\rangle$, where the symmetric product consists of n copies of $|v\rangle$, the occupation number state takes the natural form:

$$|n_1, n_n, \ldots, n_\ell, \ldots\rangle = \frac{1}{\sqrt{\prod_k n_k!}} |e_1\rangle^{n_1} \vee |e_2\rangle^{n_2} \vee \ldots \vee |e_\ell\rangle^{n_\ell} \vee \ldots\ .$$

Theorem 2.3.5.
The occupation number states are a Hilbert basis of the bosonic Fock space.

Proof 2.3.6.
If $\sum_\ell n'_\ell \neq \sum_\ell n_\ell$, the scalar product is zero by definition. Assume $\sum_\ell n'_\ell \neq \sum_\ell n_\ell$ and let $n_i = n'_j\ \forall\ i,j$:

$$\langle n'_1, n'_2, ... | n_1, n_2, ...\rangle = \frac{1}{\prod_k n_k!} \cdot (n_1! \langle e_1|e_1\rangle \cdot n_2! \langle e_2|e_2\rangle \cdot \ldots) = 1\ .$$

If there is one j with $n_j \neq n'_j$, the above scalar product becomes zero.

A Hilbert basis of $S^n(\mathcal{H})$ is given by $\{|e_{j_1}\rangle \vee ... \vee |e_{j_n}\rangle\}_{j_1 \leq ... \leq j_n}$. For all $|i_1\rangle \vee ... \vee |i_N\rangle$ it holds that:

$$|0, ..., 0, \underbrace{1}_{j_1}, 0, ..., 0, \underbrace{1}_{j_n}, 0, ...\rangle = |j_1\rangle \vee ... \vee |j_n\rangle\ .$$

□

Lemma 2.3.7.
In the occupation number Hilbert basis, the creation/annihilation operators are

characterized as follows:

$$a_j^\dagger|n_1,...,n_j,...\rangle = \sqrt{n_j+1}\ |n_1,...,n_j+1,...\rangle\ ,$$
$$a_j|n_1,...,n_j,...\rangle = \sqrt{n_j}|n_1,...,n_j-1,...\rangle\ .$$

Proof 2.3.8.

i)

$$\begin{aligned} a_j^\dagger|n_1,...,n_j,...\rangle &= \frac{1}{\sqrt{\prod_\ell n_\ell!}}(|e_1\rangle^{n_1}\vee|e_2\rangle^{n_2}\vee...\vee|\ell\rangle^{n_\ell}\vee...|e_j\rangle^{n_j+1}\vee...) \\ &= \frac{\sqrt{n_j+1}}{\sqrt{(n_j+1)!}}\frac{1}{\sqrt{\prod_{\ell\neq j} n_\ell!}}(|e_1\rangle^{n_1}\vee|e_2\rangle^{n_2}\vee...\vee|e_\ell\rangle^{n_\ell}\vee...|e_j\rangle^{n_j+1}\vee...) \\ &= \sqrt{n_j+1}|n_1,...,n_j+1,...\rangle \end{aligned}$$

ii)

$$\begin{aligned} a_j|n_1,...,n_j,...\rangle &= n_j\cdot\frac{1}{\sqrt{\prod_\ell n_\ell!}}(|e_1\rangle^{n_1}\vee|e_2\rangle^{n_2}\vee...\vee|e_\ell\rangle^{n_\ell}\vee...|e_j\rangle^{n_j-1}\vee...) \\ &= \frac{n_j}{\sqrt{n_j!}}\frac{1}{\sqrt{\prod_{\ell\neq j} n_\ell!}}(|e_1\rangle^{n_1}\vee|e_2\rangle^{n_2}\vee...\vee|\ell\rangle^{n_\ell}\vee...|e_j\rangle^{n_j-1}\vee...) \\ &= \sqrt{n_j}|n_1,...,n_j-1,...\rangle \end{aligned}$$

□

Corollary 2.3.9.
As the notation suggests, the creation/annihilation operators are adjoint to each other.

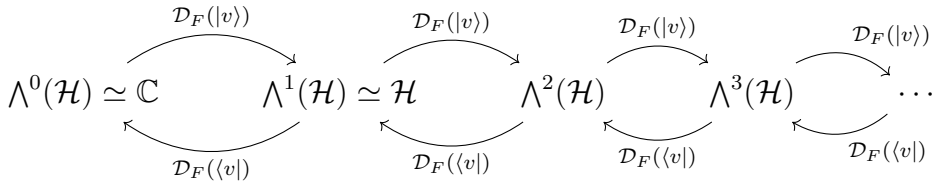
Proof 2.3.10.

$$\begin{aligned} \langle n_1',n_2',...|a_j^\dagger|n_1,n_2,...\rangle &= \sqrt{n_j+1}\ \delta_{n_j' n_j+1}\prod_{\ell\neq j}\delta_{n_j' n_\ell} = \sqrt{n_j'}\ \delta_{n_j'-1\ n_j}\prod_{\ell\neq j}\delta_{n_j' n_\ell} \\ &= (a_j\ |n_1',n_2',...\rangle, |n_1,n_2,...\rangle)\ . \end{aligned}$$

□

2.3.2. Fermions

In the case of fermions, the description is almost the same, using the fermionic Fock space representation $\mathcal{D}_F$:



As before, the **creation/annihilation operators** $c_j^\dagger / c_j$ are defined by

$$c_j^\dagger \coloneqq \mathcal{D}_F(|e_j\rangle) \qquad \text{and} \qquad c_j \coloneqq \mathcal{D}_F(\langle e_j|) \ .$$

Corollary 2.3.11.
The fermionic creation/annihilation operators satisfy the ***canonical anticommutator relations****:*

$$\{c_j^\dagger, c_k^\dagger\} = 0 \qquad \{c_j, c_k\} = 0 \qquad \{c_k, c_j^\dagger\} = \delta_{jk}\mathbb{1} \qquad \forall\ j,k\ .$$

Proof 2.3.12.
This is just a special case of theorem D.4.4. □

The different algebraic structure of $\bigwedge(\mathcal{H})$ compared to $S(\mathcal{H})$ results in a different behavior of fermions. Most notably is the Pauli exclusion principle. Due to the antisymmetry of the wedge product, $|v\rangle^n = |v\rangle \wedge \ldots \wedge |v\rangle$ is non-zero only for $n \in \{0, 1\}$.

Pauli exclusion principle

Given state $|v\rangle$, then there can be either no ore only one particle in that state, due to the antisymmetry of fermionic states.

The occupation number states are define in the same way as before. As consequence of the Pauli exclusion principle $n_l \in \{0, 1\}$, there is no need for normalization coefficients $(0! = 1! = 1)$:

$$|n_1, n_n, \ldots, n_\ell, \ldots\rangle \coloneqq (c_1^\dagger)^{n_1}(c_2^\dagger)^{n_2} \ldots (c_\ell^\dagger)^{n_\ell} \ldots |0\rangle \ .$$

The occupation number states are also a Hilbert basis for the fermionic Fock space. The proof carries over without much changes. However, the creation/annihilation operators look slightly different in the fermionic case, since the operators do not commute anymore:

$$c_j^\dagger |n_1, ..., n_j, ...\rangle = (1 - n_j)(-1)^{\sum_{i<j} n_i} \ |n_1, ..., n_j + 1, ...\rangle \ .$$
$$c_j |n_1, ..., n_j, ...\rangle = n_j (-1)^{\sum_{i<j} n_i} \ |n_1, ..., n_j - 1, ...\rangle \ .$$

2.4. Second quantization of operators

We know already, how to construct operators on tensor spaces. The explicit construction of the symmetric and wedge products, allows to extend the construction for both Fock spaces. This will lead to the second quantization of operators, that can be found in most textbooks. Before we do so however, we have to make some preparations:

Assume a tensor operator L_α of the form

$$L_\alpha = \mathbb{1} \otimes ... \otimes \underbrace{L}_{\alpha} \otimes ... \otimes \mathbb{1} \ ,$$

where $L \in \mathrm{End}(\mathcal{H})$. The sum over all postions of these operators becoms:

$$\begin{aligned}\sum_{\alpha=1}^{N} L_\alpha(|v_1\rangle \vee ... \vee |v_N\rangle) &= \sum_{\alpha=1}^{N} L_\alpha S_+ |v_1\rangle ... |v_N\rangle \\ &= L|v_1\rangle \vee ... \vee |v_N\rangle + |v_1\rangle \vee L|v_2\rangle \vee ... \vee |v_n\rangle + ... + |v_1\rangle \vee ... \vee L|v_N\rangle \ ,\end{aligned}$$

$$\begin{aligned}\sum_{\alpha=1}^{N} L_\alpha(|v_1\rangle \wedge ... \wedge |v_N\rangle) &= \sum_{\alpha=1}^{N} L_\alpha S_- |v_1\rangle ... |v_N\rangle \\ &= L|v_1\rangle \wedge ... \wedge |v_N\rangle + |v_1\rangle \wedge L|v_2\rangle \wedge ... \wedge |v_n\rangle + ... + |v_1\rangle \wedge ... \wedge L|v_N\rangle \ .\end{aligned}$$

In fact, without the construction from tensor operators, we could also have defined the meaning of $L_1 \vee L_2$ and $L_1 \wedge L_2$ accordingly. However, operators of the form $\sum_\alpha L_\alpha$ are enough for our purposes here.

2.4.1. Occupation number operator and exchange operators

Definition 2.4.1.
The operators $N_j \coloneqq a_j^\dagger a_j$ and $N_j \coloneqq c_j^\dagger c_j$ respectively, are called **Occupation number operators**.

Lemma 2.4.2.
The occupation number operator returns the number of particles in the state $|e_j\rangle$.

Proof 2.4.3.
Bosons

$$\begin{aligned}N_j|n_1, ..., n_j, ...\rangle &= \sqrt{n_j} a_j^\dagger |n_1, ..., n_j - 1, ...\rangle = \sqrt{n_j}\sqrt{n_j}|n_1, ..., n_j, ...\rangle \\ &= n_j|n_1, ..., n_j, ...\rangle\end{aligned}$$

Fermions

$$\begin{aligned}N_j|n_1, ..., n_j, ...\rangle &= n_j(-1)^{\sum_{i<j} n_j} c_j^\dagger \ |n_1, ..., n_j - 1, ...\rangle \\ &= (1 - n_j + 1)n_j(-1)^{\sum_{i<j} n_j}(-1)^{\sum_{i<j} n_j}|n_1, ..., n_j, ...\rangle \\ &= (2 - n_j)n_j|n_1, ..., n_j, ...\rangle \overset{n_j \in \{0,1\}}{=} n_j|n_1, ..., n_j, ...\rangle\end{aligned}$$

□

We define the **one-body exchange operator** σ_{mn} by $\sigma_{mn} = \sum_\alpha |e_m\rangle\langle e_n|_\alpha$, with

$$|e_m\rangle\langle e_n|_\alpha \coloneqq \mathbb{1} \otimes \ldots \otimes \underbrace{|e_m\rangle\langle e_n|}_{\alpha} \otimes \ldots \otimes \mathbb{1} \ .$$

Lemma 2.4.4.
For both bosons and fermions, the one-body exchange operator acts on the occupation number states as follows:

$$\sigma_{mn}|r_1, r_2, \ldots\rangle = a_m^\dagger a_n |r_1, r_2, \ldots\rangle$$

$$\textit{and} \quad \sigma_{mn}|r_1, r_2, \ldots\rangle = c_m^\dagger c_n |r_1, r_2, \ldots\rangle$$

Proof 2.4.5.
For bosons, the calculation is straight forward:

$$\begin{aligned}
\sigma_{mn}|r_1, r_2, ...\rangle &= \frac{1}{\prod_{j=1}^{\infty}\sqrt{r_j!}} \sum_\alpha (|e_m\rangle\langle e_n|)_\alpha \; |e_1\rangle^{r_1} \vee ... \vee |e_m\rangle^{r_m} \vee ... \vee |e_n\rangle^{r_n} \vee \ldots \\
&= \frac{r_n}{\prod_{j=1}^{\infty}\sqrt{r_j!}} \; |e_1\rangle^{r_1} \vee ... \vee |e_m\rangle^{r_m+1} \vee ... \vee |e_n\rangle^{r_n-1} \vee \ldots \\
&= r_n\sqrt{r_m+1}\frac{1}{\sqrt{r_n}}|r_1, ..., r_m+1, ..., r_n-1\rangle \\
&= \sqrt{r_m+1}\sqrt{r_n}|r_1, ..., r_m+1, ..., r_n-1\rangle \\
&= a_m^\dagger a_n |r_1, r_2, ...\rangle \; .
\end{aligned}$$

Fermions need further attention, for there is no commutation of elements. Yet, no normalization factors appear ($r_j \in \{0, 1\}$). Assume Œ $n > m$. Also let p denote the number of elements missing between $|e_m\rangle$ and $|e_n\rangle$:

$$\begin{aligned}
\sigma_{mn}|r_1, r_2, ...\rangle &= \sum_\alpha (|e_m\rangle\langle e_n|)_\alpha \; |e_1\rangle^{r_1} \wedge ... \wedge |e_m\rangle^{r_m} \wedge ... \wedge |e_n\rangle^{r_n} \wedge \ldots \\
&= (1-r_m)r_n|e_1\rangle^{r_1} \wedge ... \wedge \wedge ... \wedge |e_m\rangle \wedge \ldots \\
&= (-1)^{\sum_{m<j<n} r_j}(1-r_m)r_n|e_1\rangle^{r_1} \wedge ... \wedge |e_m\rangle^{r_m+1} \wedge ... \wedge |e_n\rangle^{r_n-1} \wedge \ldots \\
&= -(1-r_m)r_n|e_1\rangle^{r_1} \wedge ... \wedge |e_m\rangle^{r_m+1} \wedge ... \wedge |e_n\rangle^{r_n-1} \wedge \ldots \\
&= -(1-r_m)r_n|r_1, \ldots, r_m+1, \ldots, r_n-1, \ldots\rangle
\end{aligned}$$

In the second line, the coefficients are the condition, that the whole term is non-zero ($r_m = 0$ and $r_n = 1$). In the third line, the $|e_m\rangle$, that was on the n-th position has been exchanged $\sum_{m<j<n} r_j$ times, to reach the m-th position. But that means $r_m \to r_m + 1$ and $r_n \to r_n - 1$.

On the other hand, we get (using $r_m = 0$ in the first equality):

$$\begin{aligned}
c_m^\dagger c_n |r_1, r_2, ...\rangle &= r_n(-1)^{\sum_{j<m} r_j + \sum_{m<j<n} r_j} c_m^\dagger |r_1, r_2, ..., r_n-1\rangle \\
&= (1-r_m)r_n(-1)^{\sum_{j<m} r_j + \sum_{m<j<n} r_j} \\
&\qquad \cdot (-1)^{\sum_{j<m} r_j}|r_1, \ldots, r_m+1, \ldots, r_n-1, \ldots\rangle \\
&= (1-r_m)r_n(-1)^{\sum_{m<j<n} r_j}|r_1, \ldots, r_m+1, \ldots, r_n-1, \ldots\rangle
\end{aligned}$$

□

In the same way, we define the **two-body exchange operator** $\sigma_{ijk\ell}$ by $\sigma_{ijk\ell} = \sum_{\alpha\neq\beta,\beta\neq\alpha} |e_i\rangle|e_j\rangle\langle e_k|_\alpha\langle e_\ell|_\beta$ with

$$\begin{aligned}|e_i\rangle|e_j\rangle\langle e_k|_\alpha\langle e_\ell|_\beta &= \mathbb{1}\otimes ...\otimes \mathbb{1}\otimes \underbrace{|e_i\rangle\langle e_k|}_{\alpha}\otimes \mathbb{1}\otimes ...\otimes \underbrace{|e_j\rangle\langle e_\ell|}_{\beta}\otimes ...\\ &= \left(\mathbb{1}\otimes ...\otimes \mathbb{1}\otimes \underbrace{|e_i\rangle\langle e_k|}_{\alpha}\otimes \mathbb{1}\otimes ...\right)\\ &\qquad \circ \left(\mathbb{1}\otimes ...\otimes \mathbb{1}\otimes \underbrace{|e_j\rangle\langle e_\ell|}_{\beta}\otimes \mathbb{1}\otimes ...\right)\\ &= |e_i\rangle\langle e_k|_\alpha \circ |e_k\rangle\langle e_\ell|_\beta\ .\end{aligned}$$

Corollary 2.4.6.
For both bosons and fermions, the two-body exchange operator acts on the occupation number states as follows:

$$\sigma_{ijk\ell} = a_i^\dagger a_j^\dagger a_k a_\ell \quad \textit{and} \quad \sigma_{ijk\ell} = c_i^\dagger c_j^\dagger c_\ell c_k\ .$$

Proof 2.4.7.
Rewriting σ_{ijkl} yields:

$$\begin{aligned}\sum_{\alpha\neq\beta,\beta\neq\alpha} |e_i\rangle|e_j\rangle\langle e_k|_\alpha\langle e_\ell|_\beta &= \left(\sum_\alpha |e_i\rangle\langle e_k|_\alpha\right)\left(\sum_\beta |e_j\rangle\langle e_\ell|_\beta\right) - \sum_\alpha |e_i\rangle\langle e_k|_\alpha \circ |e_j\rangle\langle e_\ell|_\alpha\\ &= a_i^\dagger a_k a_j^\dagger a_\ell - \delta_{k,j} a_i^\dagger a_\ell \qquad \text{or} \qquad c_i^\dagger c_k c_j^\dagger c_\ell - \delta_{k,j} c_i^\dagger c_\ell\ .\end{aligned}$$

For Bosons, we get:

$$a_i^\dagger a_k a_j^\dagger a_\ell - \delta_{k,j} a_i^\dagger a_\ell \underbrace{=}_{[a_k, a_j^\dagger]=\delta_{j,k}\mathbb{1}} a_i^\dagger a_j^\dagger a_k a_\ell\ .$$

For Fermions, we get:

$$c_i^\dagger c_k c_j^\dagger c_\ell - \delta_{k,j} c_i^\dagger c_\ell \underbrace{=}_{\{c_k, c_j^\dagger\}=\delta_{k,j}\mathbb{1}} - c_i^\dagger c_j^\dagger c_k c_\ell = c_i^\dagger c_j^\dagger c_\ell c_k\ .$$

□

2.4.2. One-body operators

One-body operators are operators in the usual sense. That is, they are operators that act on each particle individually, without any interaction terms. The prime examples would be the kinetic energy and external potentials. These operators can be written as

$$\mathbf{T} = \sum_\alpha T_\alpha\ ,$$

where all T_α have the same operator $T \in \mathrm{End}(\mathcal{H})$ at position α. From

$$T = \mathbb{1}T\mathbb{1} = \sum_{m,n} |e_m\rangle\langle e_m|T|e_n\rangle\langle e_n| = \sum_{m,n} T_{mn}|e_m\rangle\langle e_n| \ ,$$

we see, that we can write $\mathbf{T}$ in the following form (the matrix elements of T do not depend upon α):

$$\mathbf{T} = \sum_{m,n} T_{mn} \sum_{\alpha} |e_m\rangle\langle e_n|_\alpha \coloneqq \sum_{m,n} T_{mn} \cdot \sigma_{m,n} \ .$$

In the last subsection, we have already investigated, how to express the one-body exchange operator σ_{mn} differently. We can write down the end result for the **second quantized one-body operator $\mathbf{T}$** directly:

$$\boxed{\mathbf{T} = \begin{cases} \sum\limits_{m,n} T_{mn}\ a_m^\dagger a_n & \text{for bosons} \\ \sum\limits_{m,n} T_{mn}\ c_m^\dagger c_n & \text{for fermions} \end{cases}}$$

Remark 2.4.8.
Let Q be the symmetric bilinear form and A the symplectic form from lemma D.4.1, as well as $T = \sum_{m,n} T_{mn}|e_m\rangle\langle e_n|$ be a linear operator on $\mathcal{H}$. Define the corresponding elements in the Weyl- and Clifford algebra T^W and T^C by

$$T^W = T = \sum_{m,n} T_{mn}|e_m\rangle \circ \langle e_n| \in \mathcal{W}(\mathcal{H} \oplus \mathcal{H}^*, A) \ ,$$

$$T^C = T = \sum_{m,n} T_{mn}|e_m\rangle \circ \langle e_n| \in \mathrm{Cl}(\mathcal{H} \oplus \mathcal{H}^*, Q) \ .$$

Recalling the definitions of creation and annihilation operators (def. 2.3.1 etc.), we recognize $\mathbf{T}$ as representation $\mathcal{D}_B(T^W)$ for bosons and $\mathcal{D}_F(T^C)$ for fermions.

2.4.3. Two-body operators

Besides external potential, there can be pair potentials between the particles. A prime example are coulomb potentials. To accommodate for these situations, two-body operators, that have to take into account all interaction between all particles, are defined.

The pair interaction on $\mathcal{H}^{\otimes n}$ between a particle at position α and a particle at position β can be expressed by

$$V_{\alpha,\beta} = \sum_{i,j,k,\ell} V_{ijkl} \cdot |e_i\rangle|e_j\rangle\langle e_j|_\alpha\langle e_\ell|_\beta \ .$$

As in classical mechanics, instead of summing over all disjoint pairs $\alpha < \beta$, we can sum over all pairs $\alpha \neq \beta$, and fix the double counting with the coefficient $\frac{1}{2}$:

$$\mathbf{V} = \frac{1}{2} \sum_{\alpha\neq\beta,\beta\neq\alpha} V_{\alpha,\beta} = \frac{1}{2} \sum_{\alpha\neq\beta,\beta\neq\alpha} \sum_{i,j,k,\ell} V_{ijk\ell} \cdot |e_i\rangle|e_j\rangle\langle e_j|_\alpha\langle e_\ell|_\beta$$

$$= \frac{1}{2} \sum_{i,j,k,\ell} V_{ijk\ell} \sigma_{ijk\ell} \ .$$

Thus, we get the common form of **second quantized two-body operators**:

$$\mathbf{V} = \begin{cases} \frac{1}{2} \sum\limits_{i,j,k,\ell} V_{ijk\ell} \ a_i^\dagger a_j^\dagger a_k a_\ell & \text{for bosons} \\ \frac{1}{2} \sum\limits_{i,j,k,\ell} V_{ijk\ell} \ c_i^\dagger c_j^\dagger c_\ell c_k & \text{for fermions} \end{cases}$$

Remark 2.4.9.
Similar to one-body operators, second quantized two-body operators $\mathbf{V}$ can be regarded as representations $\mathcal{D}_B(V^W)$ and $\mathcal{D}_F(V^C)$.

3
Dirac equation

The Dirac equation is an attempt to describe relativistic quantum mechanics. Despite its problems, that get resolved in quantum field theory, the Dirac theory gives rise to the description of spin in a natural way. It also suggested the existence of anti particles and it (or its field theoretic counterpart) shaped the development of relativistic quantum mechanics significantly. This chapter is again based on [Zir10].

3.1. Motivation

The Schrödinger equation is a quantum mechanical version of classical Hamiltonian mechanics. Since we already know, that special relativity generalizes classical mechanics[1], a quantum theory is expected to satisfy the energy momentum relation $E = \sqrt{p^2c^2 + m^2c^4}$. The first attempt, using the correspondences $E \leftrightarrow i\hbar\partial_t$ and $p^2 \leftrightarrow -\hbar^2\nabla^2$, suffers from the problem, that an operator

$$\sqrt{\hbar^2c^2\nabla^2 + m^2c^4}$$

does not seem to be reasonable. In a power series representation such an operator would be non-local, containing derivatives up to infinite order. An attempt to fix this problem is to use the squared energy momentum relation:

$$-\hbar^2\partial_t^2\psi = (\hbar^2c^2\nabla^2 + m^2c^4)\psi \ .$$

This equation is known as **Klein-Gordon equation**. We will not inquire this equation any further, referring to the literature instead, but mention that the Klein-Gordon equation does not give a continuity equation for a probability density. However, this was an expectation at the time the equation was developed. The problem arises due to the second order time derivative.

Motivated by this problem, Dirac sought to find a Hamilton operator H, that satisfies:

- H is linear in all derivatives,
- solutions to H satisfy the squared energy momentum relation.

The second condition has the same reason as before, finding a quantum mechanical generalization that is compatible with special relativity. The first condition, that can be regarded as motivation to overcome the problems of the Klein-Gordon equation, are motivated by the following observation: In special relativity (as well as general relativity) space and time are described as one structure, having one expect a first order partial time derivative needing first order spatial derivatives.

[1] In fact, general relativity is a further generalization. However, so far no one has been able to unify quantum mechanics and general relativity.

Guided by these principles, the following ansatz is employed:

$$H = \beta mc^2 + c^2 \sum_{j=1}^{3} \alpha_j p_j^2 \qquad \text{with} \qquad p_j = -i\hbar\partial_{x_j} \ .$$

To recover the energy momentum relation, the following algebraic relations are needed:

$$\beta^2 = 1 \qquad \beta\alpha_j + \alpha_j\beta = 0 \ , \qquad \alpha_j\alpha_k + \alpha_k\alpha_j = 2\delta_{jk} \ . \tag{3.1}$$

From these relations, it is clear that the Dirac coefficients β and α_j are no elements of any algebraic field. The **Dirac equation** then is

$$\boxed{i\hbar\partial_t\psi = H\psi = \left(mc^2\beta - i\hbar c\sum_{j=1}^{3}\alpha_j\partial_{x_j}\right)\psi \ .}$$

3.2. Spinors and Dirac equation

To find a realization for β and α_j we define new objects, derived from the Dirac coefficients:

$$\gamma^0 = \beta \ , \qquad \gamma^j = \beta\alpha_j = -\alpha_j\beta \ . \tag{3.2}$$

With the Lorentzian metric $g = \mathrm{diag}(1, -1, -1, -1)$ and the standard convention of special relativity, $x^0 = ct$, $x^j = x_j$, we obtain the new relations:

$$\boxed{\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu} \ .} \tag{3.3}$$

Proof 3.2.1.
Assuming, the relations for the Dirac coefficients are satisfied, then:

$$\gamma^0\gamma^0 + \gamma^0\gamma^0 = 2\beta^2 = 2 = 2g^{00} \ ,$$

$$\gamma^0\gamma^j + \gamma^j\gamma^0 = \beta\beta\alpha_j - \alpha_j\beta\beta = \alpha_j - \alpha_j = 0 \ ,$$

$$\gamma^i\gamma^j + \gamma^j\gamma^i = -(\alpha_i\beta^2\alpha_j + \alpha_j\beta^2\alpha_i) = -(\alpha_i\alpha_j + \alpha_j\alpha_i) = -2\delta_{ij} = 2g^{ij} \ .$$

The other direction follows from the same equations. □

Lemma 3.2.2.
The Dirac equation, written in terms of γ^μ, has the form

$$\left(\gamma^\mu\partial_\mu + i\frac{mc}{\hbar}\right)\psi = 0 \ .$$

This equation is called ***covariant Dirac equation***

Proof 3.2.3.

$$0 = \left(\gamma^\mu \frac{\partial}{\partial x_\mu} + i\frac{mc}{\hbar}\right)\psi = \left(\frac{\beta}{c}\partial_t - \beta\sum_{j=1}^{3}\alpha_j\partial_{x_j} + i\frac{mc}{\hbar}\right)\psi$$

$$\overset{\beta^2=1}{\Leftrightarrow} \qquad \partial_t\psi = -\frac{i}{\hbar}\beta mc^2\psi + c\sum_{j=1}^{3}\alpha_j\partial_{x_j}$$

$$\Leftrightarrow \qquad i\hbar\partial_t\psi = mc^2\beta\psi - i\hbar c\sum_{j=1}^{3}\alpha_j\partial_{x_j}\psi$$

□

Remark 3.2.4.
The name "covariant" comes from the fact, that the covariant Dirac equation is form invariant under rotations in the Minkowski space. We will prove this in the context of electromagnetic fields in theorem 3.3.1.

After we have established the Dirac equation in this new form, using the γ^μ-s, we can investigate the relations (3.3), that look like Clifford algebra relations. Remembering the spinor representation (see section D.3), the γ^μ-s can be understood as linear maps as well as elements of a Clifford algebra.

The spinor representation in section D.3 was constructed for the Euclidean scalar product, on the space $\bigwedge(P)$. Here, we have the Minkowski scalar product, and, as the position of the indices symbolize, covariant objects. The former will need a different polarization, while the latter has no severe impact.

Theorem 3.2.5.
Let $(\mathbb{R}^4, g)$ be the Minkowski space with g-orthonormal basis $\{e_0, \dots, e_3\}$ an choose the polarization $\mathbb{C}^4 = P \oplus P^*$ with

$$P = span_{\mathbb{C}}\left\{c_1 := \frac{1}{2}(e_0 + e_3)\ , \quad c_2 := \frac{1}{2}(e_1 - ie_2)\right\}\ ,$$

$$P^* = span_{\mathbb{C}}\left\{c_1^* := \frac{1}{2}(e_0 - e_3)\ , \quad c_2^* := \frac{1}{2}(e_1 + ie_2)\right\}\ .$$

Let $\mathcal{D}_S\colon \mathrm{Cl}(\mathbb{C}^4, g) \to \mathrm{End}(\bigwedge(P))$ be the spinor representation, then $\gamma^\mu := g^{\mu\nu}\gamma_\nu := g^{\mu\nu}\mathcal{D}_S(e_\nu)$ satisfy the relations (3.3) and have the following matrix representation:

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1}_{\mathbb{C}^2} \\ \mathbb{1}_{\mathbb{C}^2} & 0 \end{pmatrix}\ , \qquad \gamma^j = \begin{pmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{pmatrix}\ ,$$

where σ_j denote the Pauli matrices.

Proof 3.2.6.
That the relations (3.3) are satisfied is a direct consequence of the homomorphy of $\mathcal{D}_S$ and the diagonality of g. Also, we will only show the claim for γ^0, as the rest follows in exactly the same way. Before we do so, we calculate $c_1^*(c_1)$:

$$c_1^*(c_1) = 2g(c_1^*, c_1) = \frac{1}{2}(g(e_0, e_0) - g(e_1, e_1)) = \frac{1}{2}(1+1) = 1 \ .$$

The basis of $\bigwedge(P)$ is $\{1, c_1 \wedge c_2, c_1, c_2\}$, where we chose this strange order, to obtain the desired matrix representation. The actions of c_1 and c_1^* are:

$$\begin{aligned} &\mathcal{D}_S(c_1)1 = c_1 \ , && \mathcal{D}_S(c_1)c_1 = 0 \ , && \mathcal{D}_S(c_1)c_2 = c_1 \wedge c_2 \ , && \mathcal{D}_S(c_1)c_1 \wedge c_2 = 0 \ , \\ &\mathcal{D}_S(c_1^*)1 = 0 \ , && \mathcal{D}_S(c_1^*)c_1 = 1 \ , && \mathcal{D}_S(c_1^*)c_2 = 0 \ , && \mathcal{D}_S(c_1^*)c_1 \wedge c_2 = c_2 \ . \end{aligned}$$

Since $e_0 = c_1 + c_1^*$ and $\mathcal{D}_S$ is also linear, we can read of the action of e_0:

$$\gamma_0 1 = \mathcal{D}_s(e_0)1 = c_1 \ , \quad \gamma_0 c_1 = \mathcal{D}_s(e_0)c_1 = 1 \ , \quad \gamma_0 c_2 = \mathcal{D}_s(e_0)c_2 = c_1 \wedge c_2 \ ,$$

$$\gamma_0 c_1 \wedge c_2 = \mathcal{D}_s(e_0)c_1 \wedge c_2 = c_2 \ .$$

Thus the matrix representation of $\gamma^0 = g^{00}\gamma_0 = \gamma_0$ choosing

$$\{1, c_1 \wedge c_2, c_1, c_2\} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

is

$$\gamma^0 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \mathbb{1}_{\mathbb{C}^2} \\ \mathbb{1}_{\mathbb{C}^2} & 0 \end{pmatrix} \ .$$

The other γ^μ-s follow in the same way. □

Remark 3.2.7.
These matrix representations of the γ^μ-s are called **Dirac matrices**, in the so called **Weyl representation**. There are other matrix representations of the Dirac matrices, for example the Dirac representation. Define the matrices

$$A = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{1}_2 & -\mathbb{1}_2 \\ \mathbb{1}_2 & \mathbb{1}_2 \end{pmatrix} \qquad A^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{1}_2 & \mathbb{1}_2 \\ -\mathbb{1}_2 & \mathbb{1}_2 \end{pmatrix} \ .$$

The Dirac representation is given by $\mathcal{D}(e_\mu) = A^{-1} \circ \mathcal{D}_S(e_\mu) \circ A$, resulting in the Dirac matrices in Dirac representation:

$$\gamma^0 = \begin{pmatrix} \mathbb{1}_{\mathbb{C}^2} & 0 \\ 0 & -\mathbb{1}_{\mathbb{C}^2} \end{pmatrix} \ , \qquad \gamma^j = \begin{pmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{pmatrix} \ .$$

Whatever the explicit representations of the Dirac matrices, they have in common to act on $\bigwedge(P)$ (or $\bigwedge(P^*)$). Hence $\psi \in \Gamma(\mathbb{R}^4, \bigwedge(P))$, that is, $\psi \colon \mathbb{R}^4 \to \bigwedge(P)$.

Definition 3.2.8.
The states ψ, that the Dirac operator acts on, are called **spinor fields**. The representation used for the Dirac matrices gives the kind of spinors, e.g. Dirac spinors and Weyl spinors.

For the rest of the chapter, we will consider Dirac spinors, if not specified differently. Furthermore, we denote the Dirac representation also by $\mathcal{D}_S$.

Remark 3.2.9.
For Dirac spinors, the matrix representations of β and α_j are:

$$\beta = \begin{pmatrix} \mathbb{1}_{\mathbb{C}^2} & 0 \\ 0 & -\mathbb{1}_{\mathbb{C}^2} \end{pmatrix} , \qquad \alpha_j = \begin{pmatrix} 0 & \sigma_j \\ \sigma_j & 0 \end{pmatrix} .$$

To find the transformation behavior of spinors, we consider the general case of maps $f\colon X \to Y$, that form a vector space Y^X.[2] Let G be a group with group actions $\mathcal{D}_X\colon G \times X \to X$ and $\mathcal{D}_Y\colon G \times Y \to Y$. A representation of G on Y^X can be defined by

$$\mathcal{D}\colon G \to \mathrm{GL}(Y^X) , \quad \mathcal{D}(g)f = \mathcal{D}_Y(g) \circ f \circ \mathcal{D}_X(g)^{-1} .$$

As group representation $\mathcal{D}$ defines an action $\mathcal{D}\colon G \times Y^X \to Y^X$. Then one can fix an element $g \in G$ and observe, that the following diagram commutes:

$$\begin{array}{ccc} X & \xrightarrow{\quad f \quad} & Y \\ {\scriptstyle \mathcal{D}_X(g)}\downarrow & & \downarrow{\scriptstyle \mathcal{D}_Y(g)} \\ X & \xrightarrow[\quad \mathcal{D}(g)f \quad]{} & Y \end{array}$$

Example 3.2.10.
Let $R \in \mathrm{SO}(3)$ be a rotation and $v\colon \mathbb{R}^3 \to \mathbb{R}^3$ a vector field. Rotations act naturally on $\mathbb{R}^3$. Vector fields transform under rotations as

$$(R \circ v)(x) = Rv(R^{-1}x) .$$

To adopt the general construction for spionor fields, we recall the spinor representation $\mathcal{D}_S$ from section D.3, observing that $\mathrm{Spin}(\mathbb{R}^4, g) \subset \mathrm{Cl}(\mathbb{C}^4, g)$, and the homomorphism $\rho\colon \mathrm{Spin}(V, Q) \to \mathrm{SO}(V, Q)$ from subsection D.5.2. Since $\mathrm{SO}(V, Q) \subset \mathrm{GL}(V)$, ρ is also a representation.

Definition 3.2.11.
Spinors transform under the spin group $\mathrm{Spin}(\mathbb{R}^4, g)$ as

$$(\mathcal{D}(a)\psi)(v) = \mathcal{D}_S(a)\psi(\rho(a)^{-1}v) , \qquad \forall\ a \in \mathrm{Spin}(\mathbb{R}^4, g) .$$

[2] Additional structure for Y is needed to define a vectors space structure on Y^X.

3.3. Non-relativistic reduction and spin

There does exist a reduction theory for the Dirac equation, that allows to take higher then linear order effects into account, called Foldy–Wouthuysen transformation. Here we will only consider the first order approximations derived by physical reasoning, as they suffice to result in the Pauli equation. The focus here is to give meaning to the Dirac equation, and how spin arises in a natural way form the theory.

3.3.1. Dirac equation with electro magnetic fields and covariance

The Faraday form $F \in \Omega^2(M_4)$ on the Minkowski space $M_4 = (\mathbb{R}^4, g)$, better known as antisymmetric field strength tensor, is generated by a potential $\mathcal{A} \in \Omega^1(M_4)$:

$$F = d\mathcal{A} \qquad \Rightarrow \qquad F_{\mu\nu} = \partial_\mu \mathcal{A}_\nu - \partial_\mu \mathcal{A}_\nu \ .$$

The property $d^2 = 0$ gives rise to the gauge invariance under closed 1-forms, that are also exact on M_4. This means, that for all $\chi \in C^\infty(M)$ and all $F = d\mathcal{A}$ it holds that:

$$\mathcal{A}' \coloneqq \mathcal{A} + d\chi \qquad \Rightarrow \qquad F = d\mathcal{A}' \ .$$

The choice of a χ is the choice of a gauge. The theory of electromagnetism is gauge invariant. From the Schrödinger equation with non-relativistic electromagnetic fields, we know the gauge transformation $\psi \to e^{\frac{i}{\hbar}\chi}\psi$. The Dirac equation is supposed to be gauge invariant. The principle of minimal substitution is the substitution $\partial_\mu \to \partial'_\mu$, such that the equation with ∂'_μ becomes gauge invariant. We find that

$$\partial_\mu \to \left(\partial_\mu - \frac{ie}{\hbar}\mathcal{A}_\mu\right)$$

satisfies our needs. The Dirac equation with electro magnetic fields is

$$\boxed{\gamma^\mu \left(\partial_{x_\mu} - \frac{ie}{\hbar}\mathcal{A}_\mu\right)\psi + i\frac{mc}{\hbar}\psi = 0 \ .}$$

Theorem 3.3.1.
Let ψ be a solution of the covariant Dirac equation and $a \in Spin(\mathbb{R}^4, g)$. Then $\mathcal{D}(a)\psi$ is a solution of the covariant Dirac equation

$$\gamma^\mu \left(\partial_{x_\mu} - \frac{ie}{\hbar}(\mathcal{D}(a)\mathcal{A})_\mu\right)\ \mathcal{D}(a)\psi + i\frac{mc}{\hbar}\mathcal{D}(a)\psi = 0 \ ,$$

with gauge transformed potential:

$$(\mathcal{D}(a)\mathcal{A})_\nu \coloneqq \mathcal{A}_\mu(\rho(a)^{-1})\left(\rho(a)^{-1}\right)^\mu{}_\nu$$

Remark 3.3.2.
The gauge transformation law is just the natural transformation of the 1-form $\mathcal{A}$,

i.e.

$$\begin{aligned}\mathcal{D}(a)\mathcal{A} &= \mathcal{D}(a)\left(\mathcal{A}_\mu(x)dx^\mu\right)(x) = \mathcal{A}_\mu(\rho(a)^{-1}x)d(\rho(a)^{-1}x)^\mu \\ &= \mathcal{A}_\mu(\rho(a)^{-1})\left(\rho(a)^{-1}\right)^\mu{}_\nu\, dx^\nu \ .\end{aligned}$$

Proof 3.3.3.
This proof follows [Zir10, section 2.12].

First, we recall the definition of $\rho(a)$ as $\rho(a)x = axa^{-1}$, where the left side is in the spin group. Thus, with γ_μ as representation of e_μ in the spinor representation, we find:

$$\begin{aligned} & ae_\mu a^{-1} = \rho(a)e_\mu = (\rho(a))^\nu{}_\mu\, e_\nu \\ \Rightarrow \qquad & ag^{\tau\mu}e_\mu a^{-1} = ae^\tau a^{-1} = g^{\tau\mu}(\rho(a))^\nu{}_\mu\, e_\nu = (\rho(a))_\nu{}^\tau\, e^\nu \ .\end{aligned}$$

Since $\rho\colon \mathrm{Spin}(\mathbb{R}^4, g) \to \mathrm{SO}(\mathbb{R}^4, g)$ it holds that $\rho(a)^T = \rho(a)^{-1}$, which looks in Ricci calculus as follows:

$$\begin{aligned} & (\rho(a))_\nu{}^\tau = ((\rho(a))^\tau{}_\nu)^T = ((\rho(a))^\tau{}_\nu)^{-1} = (\rho(a)^{-1})^\tau{}_\nu \\ \Rightarrow \qquad & ae^\tau a^{-1} = (\rho(a)^{-1})^\tau{}_\nu e^\nu \ .\end{aligned}$$

Applying the spinor representation, which by definition is a homomorphism, on both sides, yields:

$$\begin{aligned} & \mathcal{D}_S(a)\gamma^\tau\mathcal{D}_S(a)^{-1} = (\rho(a)^{-1})^\tau{}_\nu\gamma^\nu \\ \Leftrightarrow \qquad & \mathcal{D}_S(a)\gamma^\tau = (\rho(a)^{-1})^\tau{}_\nu\gamma^\nu\ \mathcal{D}_S(a) \ .\end{aligned}$$

Starting from the Dirac equation

$$\gamma^\mu\left(\partial_{x_\mu} - \frac{ie}{\hbar}\mathcal{A}_\mu(x)\right)\psi(x) + i\frac{mc}{\hbar}\psi(x) = 0 \ ,$$

we multiply with $\mathcal{D}_S(a)$ from the left, and use our previous findings:

$$(\rho(a)^{-1})^\mu{}_\nu\gamma^\nu\left(\partial_{x_\mu} - \frac{ie}{\hbar}\mathcal{A}_\mu(x)\right)\mathcal{D}_S(a)\psi(x) + i\frac{mc}{\hbar}\mathcal{D}_S(a)\psi(x) = 0 \ .$$

Next we transform the arguments, using

$$\partial_{y_\mu} = \frac{\partial}{\partial y^\mu} = \frac{\partial x(y)^\nu}{\partial y^\mu}\frac{\partial}{\partial x^\nu} = (\rho(a))^\nu{}_\mu\frac{\partial}{\partial x^\nu} \ ,$$

where $y = (\rho(a)^{-1}x$ and $x(y) = \rho(a)y$:

$$\begin{aligned} 0 &= (\rho(a)^{-1})^\mu{}_\nu\gamma^\nu\left((\rho(a))^\lambda{}_\mu\partial_{x_\lambda} - \frac{ie}{\hbar}\mathcal{A}_\mu(\rho(a)^{-1}x)\right)\mathcal{D}_S(a)\psi(\rho(a)^{-1}x) \\ &\quad + i\frac{mc}{\hbar}\mathcal{D}_S(a)\psi(\rho(a)^{-1}x) \\ &= (\rho(a)^{-1})^\mu{}_\nu(\rho(a))^\lambda{}_\mu\gamma^\nu\left(\partial_{x_\lambda} - \frac{ie}{\hbar}\mathcal{A}_\mu(\rho(a)^{-1}x)\ (\rho(a)^{-1})^\mu{}_\nu\right)\mathcal{D}(a)\psi\end{aligned}$$

$$
\begin{aligned}
&\quad + i\frac{mc}{\hbar}\mathcal{D}(a)\psi \\
&= \delta^\lambda_\nu \left(\partial_{x_\lambda} - \frac{ie}{\hbar}(\mathcal{D}(A)\mathcal{A})_\nu\right)\mathcal{D}(a)\psi + i\frac{mc}{\hbar}\mathcal{D}(a)\psi \\
&= \left(\partial_{x_\nu} - \frac{ie}{\hbar}(\mathcal{D}(A)\mathcal{A})_\nu\right)\mathcal{D}(a)\psi + i\frac{mc}{\hbar}\mathcal{D}(a)\psi \ .
\end{aligned}
$$

□

Corollary 3.3.4.
If ψ is a solution of the free Dirac equation, so is $\mathcal{D}(a)\psi$.

3.3.2. Continuity equation

Writing out the Dirac equation with electromagnetic fields, in terms of β and α_ℓ yields:

$$\left(\partial_t + \frac{ie}{\hbar}\Phi\right)\psi + c\sum_\ell \left(\partial_{x_\ell} - \frac{ie}{\hbar}\mathcal{A}_\ell\right)\alpha_\ell\psi + i\frac{mc^2}{\hbar}\beta\psi = 0 \ . \tag{3.4}$$

For the spinors, we have chosen a basis of $\bigwedge(P) \simeq \mathbb{C}^4$, which allows to construct a hermitian scalar product $(\cdot,\cdot)\colon \mathbb{C}^4 \times \mathbb{C}^4 \to \mathbb{C}$. We assume, that this hermitian scalar product has a meaning beyond its construction here.

The scalar product allows to define dual spinors $\psi^\dagger \in (\mathbb{C}^4)^*$ by $\psi^\dagger = (\overline{\psi}, \cdot)$. The hermitian adjoint matrices $\beta^\dagger$ and $\alpha_l^\dagger$ are defined as always. A short calculation reveals β and α_ℓ to be hermitian. The adjoint equation thus is

$$\left(\partial_t - \frac{ie}{\hbar}\Phi\right)\psi^\dagger + c\sum_\ell \left(\partial_{x_\ell} + \frac{ie}{\hbar}\mathcal{A}_\ell\right)\psi^\dagger\alpha_\ell + -i\frac{mc^2}{\hbar}\psi^\dagger\beta = 0 \ . \tag{3.5}$$

Calculating the term $\psi^\dagger(3.4) + (3.5)\psi$, results in a continuity equation:

$$\partial_t \underbrace{\psi^\dagger\psi}_{=\rho} + \sum_\ell \partial_{x_\ell} \underbrace{c\psi^\dagger\alpha_\ell\psi}_{=j_\ell} = 0 \qquad \Rightarrow \qquad \partial_t\rho + \nabla\cdot\vec{j} = 0 \ .$$

3.3.3. Non-relativistic limit and Pauli equation

As is the case for any generalization of a theory in physics, the former should be a limit case. In this case, the Schrödinger equation has to be recovered from the Dirac equation in the non-relativistic limit. Since we may choose $\mathcal{A} = 0$ at any point, we will use this opportunity to discover a first order correction for the Schrödinger equation, the so called Pauli equation.

Writing the Dirac equation with electro magnetic fields in matrix form, we get

$$0 = \begin{pmatrix} (\frac{i}{\hbar}mc + \frac{1}{c}\partial_t - \frac{ie}{\hbar}\mathcal{A}_0)\mathbb{1}_2 & \sum_j(\sigma_j\partial_{x_j} - \frac{ie}{\hbar}\mathcal{A}_j) \\ -\sum_j \sigma_j(\partial_{x_j} - \frac{ie}{\hbar}\mathcal{A}_j) & (\frac{i}{\hbar}mc - \frac{1}{c}\partial_t + \frac{ie}{\hbar}\mathcal{A}_0)\mathbb{1}_2 \end{pmatrix}\begin{pmatrix}\psi_+ \\ \psi_-\end{pmatrix} \qquad \text{with} \qquad \psi_\pm = \begin{pmatrix}\psi_{\pm,\uparrow} \\ \psi_{\pm,\downarrow}\end{pmatrix} \ .$$

To shorten the notation, we write

$$0 = \begin{pmatrix} \widehat{A} & \widehat{B} \\ \widehat{C} & \widehat{D} \end{pmatrix} \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} \qquad \Rightarrow \qquad \begin{aligned} 0 &= \widehat{A}\psi_+ + \widehat{B}\psi_- \\ 0 &= \widehat{C}\psi_+ + \widehat{D}\psi_- \end{aligned} \ .$$

Assuming that $\widehat{D}$ is invertible, or an inversion can be approximated in the non-relativistic limit, we may write

$$\psi_- = -\widehat{D}^{-1}\widehat{C}\psi_+ \qquad \Rightarrow \qquad (A - \widehat{B}\widehat{D}^{-1}\widehat{C})\psi_+ = 0 \ . \tag{3.6}$$

The non-relativistic limit for weak fields is characterized by the fact, that $mc^2 \gg \frac{e}{\hbar}\mathcal{A}_0$. For that reason, the main contribution to the time evolution is determined by

$$0 = \begin{pmatrix} (\frac{i}{\hbar}mc + \frac{1}{c}\partial_t)\mathbb{1}_2 & 0 \\ 0 & (\frac{i}{\hbar}mc - \frac{1}{c}\partial_t)\mathbb{1}_2 \end{pmatrix} \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} \qquad \Rightarrow \qquad \psi^{(0)} \sim e^{-i\frac{mc^2}{\hbar}t} \ .$$

With this contribution, we can approximate $\widehat{D}^{-1}$. Using $-\frac{1}{c}\partial_t\psi \approx \frac{imc^2}{\hbar}\psi$, we find

$$\widehat{D}\psi_+ \approx \frac{2imc^2}{\hbar}\psi_+ \qquad \longrightarrow \qquad \widehat{D}^{-1} \approx \frac{\hbar}{2imc^2}\mathbb{1}_2 \ .$$

Using $\{\sigma_j, \sigma_k\} = 2\delta_{jk}\mathbb{1}$ and $[\sigma_j, \sigma_k] = 2i\sum_\ell \varepsilon_{jk\ell}\sigma_\ell$, the operator $\widehat{B}\widehat{C}$ can be calculated:

$$\begin{aligned}
\widehat{B}\widehat{C}\psi &= \left(\sum_j \sigma_j \left(\partial_{x_j} - \frac{ie}{\hbar}\mathcal{A}_j\right)\right)\left(\sum_k \sigma_k \left(\partial_{x_k} - \frac{ie}{\hbar}\mathcal{A}_k\right)\right)\psi \\
&= \sum_{jk}\left(\partial_{x_j}\sigma_j\sigma_k\left(\partial_{x_k} - \frac{ie}{\hbar}\mathcal{A}_k\right) - \frac{ie}{\hbar}\mathcal{A}_j\sigma_j\sigma_k\left(\partial_{x_k} - \frac{ie}{\hbar}\mathcal{A}_k\right)\right)\psi \\
&= \sum_{jk}\sigma_j\sigma_k\left(\partial_{x_j}\partial_{x_k} - \frac{ie}{\hbar}\partial_{x_j}\mathcal{A}_k - \frac{ie}{\hbar}\mathcal{A}_j\partial_{x_k} + \left(\frac{ie}{\hbar}\right)^2\mathcal{A}_j\mathcal{A}_k\right)\psi \\
&= 2\sum_j\left(\partial_{x_j}\partial_{x_j} - \frac{ie}{\hbar}\partial_{x_j}\mathcal{A}_j - \frac{ie}{\hbar}\mathcal{A}_j\partial_{x_j} + \left(\frac{ie}{\hbar}\right)^2\mathcal{A}_j^2\right)\psi \\
&\qquad - \sum_{j\neq k, k\neq j}\sigma_j\sigma_k\left(\frac{ie}{\hbar}\partial_{x_j}\mathcal{A}_k + \frac{ie}{\hbar}\mathcal{A}_j\partial_{x_k}\right)\psi \\
&= 2\sum_j\left(\partial_{x_j} - \frac{ie}{\hbar}\mathcal{A}_j\right)^2\psi - \frac{ie}{\hbar}\sum_{j\neq k, k\neq j}\sigma_j\sigma_k\left(\partial_{x_j}\mathcal{A}_k\right)\psi \\
&\qquad \underbrace{- \sum_{j\neq k, k\neq j}\sigma_j\sigma_k\left(\frac{ie}{\hbar}\mathcal{A}_k\partial_{x_j} + \frac{ie}{\hbar}\mathcal{A}_j\partial_{x_k}\right)\psi}_{=0} \\
&= 2\sum_j\left(\partial_{x_j} - \frac{ie}{\hbar}\mathcal{A}_j\right)^2\psi + \frac{e}{\hbar}\sum_{j,k,\ell}\varepsilon_{jk\ell}\sigma_\ell\left(\partial_{x_j}\mathcal{A}_k\right)\psi \\
&= \left(2\sum_j\left(\partial_{x_j} - \frac{e}{\hbar}\mathcal{A}_j\right)^2\mathbb{1} + 2\sum_\ell \sigma_\ell B_\ell\right)\psi \ .
\end{aligned}$$

Since $\widehat{D}^{-1}$ is proportional to $\mathbb{1}_2$, it commutes with $\widehat{B}$. Inserting the operators $\widehat{D}^{-1}$ and $\widehat{B}\widehat{C}$ in (3.6) results in the **Pauli equation** with energy shift mc^2:

$$\boxed{i\hbar\partial_t\psi_+ = (mc^2 + e\Phi)\psi_+ - \frac{\hbar^2}{2m}\sum_j\left(\partial_{x_j} - \frac{ie}{\hbar}\mathcal{A}_j\right)^2\psi_+ - \frac{e\hbar}{2m}\sum_\ell \sigma_\ell B_\ell\psi_+ \ .}$$

Here, B_ℓ are the components of the magnetic field and $\Phi = -c\mathcal{A}_0$ is the electrical potential.

3.3.4. Spin

The Pauli equation has an additional contribution to the energy, of the form $\frac{e\hbar}{2m}\sum_\ell \sigma_\ell B_\ell$. In classical mechanics, a charged particle, moving on a circular orbit, creates a magnetic moment of the form $\vec{\mu} = \frac{e}{2m}\vec{L}$. In a magnetic field, the potential energy is $E = -\vec{B}\cdot\vec{\mu}$. For that reason, σ_j could be considered as component of some intrinsic rotation. However, the commutator relations have an additional factor 2, that needs to be fixed, in order to understand the energy contribution as something related to angular momentum (compare theorem B.3.2). This has the following consequences:

- There is an intrinsic pseudo rotation[3], called **Spin** with the operator $S_j = \frac{\hbar}{2}\sigma_j$.
- The additional factor $\frac{1}{2}$ needs a gyromagnetic factor $g = 2$, such that $\mu_j = g\frac{e}{2m}S_j$.

That spin is no classical rotation can be understand, apart from its mathematical origin, from the gyromagnetic factor $g = 2$, that does not fit classical electrodynamics.

For Dirac spinors, the spin operator is defined by

$$\mathcal{S}_j = \frac{\hbar}{2}\begin{pmatrix}\sigma_j & 0\\ 0 & \sigma_j\end{pmatrix} \ .$$

The appearance of spin $\frac{1}{2}$ shows, that the Dirac equation describes only spin $\frac{1}{2}$ particles.

3.4. Total angular momentum

The Dirac equation gives rise to a new quantity, called spin. In the non-relativistic reduction, new terms appeared, that we interpreted as some kind of intrinsic angular momentum. In fact, this point of view is not only an assumption by comparison to classical physics, but has a mathematical foundation. The total angular momentum is the generator of spatial rotations of spinors. To see, how such a generator acts on spinors, we choose $X \in \mathfrak{spin}(\mathbb{R}^4, g)$, such that $\exp(tX) \in \mathrm{Spin}(\mathbb{R}^3) \subset \mathrm{Spin}(\mathbb{R}^4, g)$.

Lemma 3.4.1.
Let $J_\ell = -\sum_{j,k}\varepsilon_{\ell jk}e_j \otimes \vartheta e_k$ for $\ell, j, k \in \{1,2,3\}$, where $\{\vartheta_\ell\}$ is the dual basis of

[3] It cannot be a proper rotation, since the Pauli matrices generate the $SU(2)$ and not $SO(3)$.

$\{e_\ell\}$. *Then it holds that*

$$\tau(\Sigma_\ell) = 2iJ_\ell \ , \qquad \text{where } \mathcal{D}_S(\Sigma_\ell) = \begin{pmatrix} \sigma_\ell & 0 \\ 0 & \sigma_\ell \end{pmatrix} \ .$$

Proof 3.4.2.
Using $[\sigma_j, \sigma_k] = 2i \sum_\ell \varepsilon_{jk\ell}\sigma_\ell$, we find

$$\varepsilon_{\ell jk}\mathcal{D}_S(\Sigma_\ell) = \frac{i}{2}[\gamma_j, \gamma_k] \ .$$

Hence, in the Clifford algebra it holds that

$$\Sigma_\ell = \frac{i}{2}\varepsilon_{\ell jk}[e_j, e_k] \ .$$

From proof D.5.10, we read of ($Q(e_i, e_j) = -\delta_{ij}$), that $[[e_j, e_k], e_n] = -4(\delta_{kn}e_j - \delta_{jn}e_k)$. Plugging in yields:

$$\tau(\Sigma_\ell)e_n = [\Sigma_\ell, e_n] = -2i(\delta_{kn}e_j - \delta_{jn}e_k) \ , \quad \text{for } \varepsilon_{\ell jk} = 1 \ .$$

A direct calculation shows, that

$$J_\ell e_n = -\delta_{kn}e_j + \delta_{jn}e_k \ , \quad \text{for } \varepsilon_{\ell jk} = 1 \ .$$

Thus:

$$\tau(\Sigma_\ell) = 2iJ_\ell \ .$$

□

We have already met the representation of elements $g \in \mathrm{Spin}(\mathbb{R}^4, g)$ on spinors:

$$(\mathcal{D}(g)\psi)(v) = \mathcal{D}_S(g)\psi(\rho(g)^{-1}v) \ .$$

To find a representation for $X \in \mathfrak{spin}(\mathbb{R}^4, g) \simeq \mathfrak{so}(\mathbb{R}^4, g)$, we use theorem D.1.14.

Theorem 3.4.3.
Let $X \in \mathfrak{spin}(\mathbb{R}^4, g)$, such that $\mathfrak{so}(3) \ni \tau(X) = \sum_\ell c_\ell J_\ell$, where J_ℓ are the generators from the previous lemma, then it holds that

$$\left.\frac{d}{dt}\right|_{t=0} \mathcal{D}(e^{tX})\psi = \left.\frac{d}{dt}\right|_{t=0} \mathcal{D}_S(e^{tX}) \circ \psi \circ \rho(e^{-tX}) = -\frac{i}{\hbar}\sum_\ell c_\ell(L_\ell + \mathcal{S}_\ell)\psi \ .$$

Proof 3.4.4.
The second equation is a generalization of corollary D.1.18:

$$\left.\frac{d}{dt}\right|_{t=0} \psi \circ \rho(e^{-tX}) = D\psi \circ \left.\frac{d}{dt}\right|_{t=0} \rho(e^{-tX}) \ .$$

Using $\rho(e^{-tX})v = e^{-t\tau(X)}v$ we find:[4]

$$\left.\frac{d}{dt}\right|_{t=0} \psi \circ \rho(e^{-tX}) = -D\psi \circ \tau(X) \ .$$

By assumption, we have $\tau(X) = \sum_\ell c_\ell J_\ell$. For the same reason as in the proof of that corollary ($D\psi = \sum_j \partial_j \psi dx^j$), we get

$$\left.\frac{d}{dt}\right|_{t=0} \psi \circ \rho(e^{-tX}) = \frac{1}{i\hbar}\sum_\ell c_\ell L_\ell \psi = -\frac{i}{\hbar}\sum_\ell c_\ell L_\ell \psi \ ,$$

with $L_\ell = -i\hbar \sum_{j,k} \varepsilon_{\ell jk} x_j \partial_k$. Using, that $\mathcal{D}_S$ is an algebra representation, and thus $\mathcal{D}_S(e^{tX}) = e^{t\mathcal{D}_S(X)}$:

$$\left.\frac{d}{dt}\right|_{t=0} \mathcal{D}_S(e^{tX}) \circ \psi = \left.\frac{d}{dt}\right|_{t=0} e^{t\mathcal{D}_S(X)} \circ \psi = \mathcal{D}_S(X) \circ \psi \ .$$

Since $\tau(X) = \sum_\ell c_\ell J_\ell$, which is equivalent to $X = \sum_\ell c_\ell \tau^{-1}(J_\ell)$. The last lemma tells us, that $\tau^{-1}(J_\ell) = -\frac{i}{2}\Sigma_\ell$. Plugging in, and using the spin operator $\mathcal{S}_\ell = \frac{\hbar}{2}\mathcal{D}_S(\Sigma_\ell)$ yields:

$$\left.\frac{d}{dt}\right|_{t=0} \mathcal{D}_S(e^{tX}) \circ \psi = \mathcal{D}_S(X) \circ \psi = -\frac{i}{\hbar}\sum_\ell \mathcal{S}_\ell \circ \psi \ .$$

In total, using the product and chain rule, we get

$$\left.\frac{d}{dt}\right|_{t=0} \mathcal{D}_S(e^{tX}) \circ \psi \circ \rho(e^{-tX})^{-1} = -\frac{i}{\hbar}\sum_\ell c_\ell (L_\ell + \mathcal{S}_\ell) \ .$$

□

As a result of the last theorem, we know that $L_\ell + \mathcal{S}_\ell$ generate rotations of spinors. We define:

Definition 3.4.5.
The **total angular momentum operator** is defined by $J_\ell = L_\ell + \mathcal{S}_\ell$.

3.5. Problems of the Dirac theory

Although the Dirac equation looks promising at first sight, the problems of the Klein-Gordon equation still exist.

[4] Similarly we could have calculated $\left.\frac{d}{dt}\right|_{t=0} \rho(e^{-tX})v = \left.\frac{d}{dt}\right|_{t=0} e^{-tX}ve^{tX} = -Xv + vX = -[X, v] = -\tau(X)v$.

3.5.1. Solution of the free Dirac equation

To solve the free Dirac equation (for β and α_j), the plane wave ansatz $\psi(t,\vec{x}) = e^{-\frac{i}{\hbar}(E_p t - \vec{p}\cdot\vec{x})}u$ with $u \in \mathbb{C}^4$ is used. This results in the following eigen value equation:

$$E_p \begin{pmatrix} u_+ \\ u_- \end{pmatrix} = \begin{pmatrix} mc^2\mathbb{1}_2 & c\sum_j \sigma_j p_j \\ c\sum_j \sigma_j p_j & -mc^2\mathbb{1}_2 \end{pmatrix} \begin{pmatrix} u_+ \\ u_- \end{pmatrix} .$$

In the case of zero momentum, the Dirac Hamiltonian has the two eigen values $E_0 = \pm mc^2$. In general, it holds that (using the shorthand $\sum_j \sigma_j p_j = \vec{\sigma}\cdot\vec{p}$):

$$\begin{aligned} (E_p - mc^2)u_+ - c(\vec{\sigma}\cdot\vec{p})u_- &= 0 \\ (E_p + mc^2)u_- - c(\vec{\sigma}\cdot\vec{p})u_+ &= 0 . \end{aligned}$$

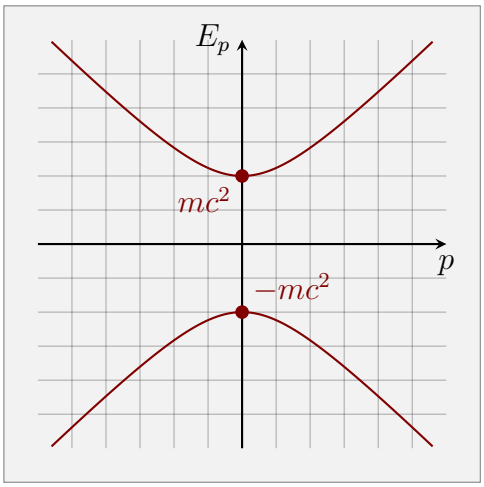


Figure 3.1.: Energy spectrum of the free Dirac Hamiltonian.

Plugging the second equation, which reads $u_- = \frac{c(\vec{\sigma}\cdot\vec{p})}{E_p+mc^2}u_+$, into the first equation results in

$$(E_p - mc^2)(E_p + mc^2)u_+ - c^2(\vec{\sigma}\cdot\vec{p})^2 u_+ = 0 .$$

A matrix calculation shows $(\vec{\sigma}\cdot\vec{p})^2 = p^2$, such that

$$E_p = \pm\sqrt{p^2c^2 + m^2c^4} .$$

We see, that the free Hamiltonian is neither bound from above nor from below. Unbound operators are no rare thing in quantum mechanics. Yet, having no lower bound poses a drastic problem. The absence of a lower bound means, that there is no ground state. This could lead to infinite transition in lower states, allowing to pull an infinite energy out of nothing, in the presence of strong electromagnetic fields, which certainly contradicts all observations.

3.5.2. Conclusion

Further investigation of the Dirac theory that we have introduced so far, reveals, that while there is a continuity equation, scattering allows for larger than unity reflections, hinting at two adaptions to be made. The continuity equation does not hold for probability densities, but for charge densities and the Dirac theory is not a single particle theory, but a many particle theory. This attempt, a modern version of the famous concept of the “Dirac sea”, leads to quantum field theory.

4

Stable second quantization for the Dirac theory

As mentioned in section 3.5, the Dirac theory still poses serious problems. As hinted at in the conclusion, the modern interpretation of the Dirac theory is a many particle formulation that allows to fix the negative spectrum and will result in a field theory. To arrive at the modern formulation from the perspective of a many particle theory, this section follows [Zir10, Sections 3.6.3 ff.] and [Zir14].

4.1. The Dirac sea and normal ordering

To fix the problem of an unbound negative spectrum, Dirac utilized the Pauli exclusion principle, and effectively reduced the Dirac theory to fermions. His idea was to interpret the vacuum as "sea" of electrons, i.e. all negative energy states being occupied by an electron. Thus, there can be no transitions of electrons into the negative spectrum, since all these states are Pauli blocked. Still, there is the possibility for the sea electrons to be promoted to real electrons, leaving a hole in the sea, that would be perceived as negative electron, today better known as positron.

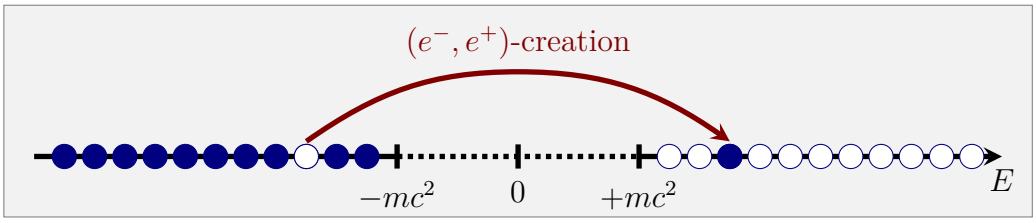


Figure 4.1.: Creation of a particle-hole pair in the Dirac sea picture. (See [Zir10, 3.6.3 The idea of hole theory.])

To sketch, how this idea can be realized, we pass to the simplest case, i.e. $\mathcal{H} \simeq \mathbb{C}$ and thus $H = h \in \mathbb{R}$. Then there is one pair of creation/annihilation operators $(c^\dagger, c)$ and $\mathbf{H} = hc^\dagger c$. If $h < 0$, the operator is negative and the naive vacuum $|0\rangle$ has to be changed, i.e. $|\text{vac}\rangle = c^\dagger|0\rangle$. Also the second quantized operator has to be adapted to

$$\widehat{\mathbf{H}} = h(c^\dagger c - 1) \ ,$$

such that $\widehat{\mathbf{H}}|\text{vac}\rangle = 0$.

In modern physics, the concept of the Dirac theory has been replaced by QFT. There might be more elaborate properties of the Dirac see, that have not been validated by experiment. Yet even in the early days of the theory, it was criticized several times. It is a great leap of faith to assume an infinite amount of charge and energy to define the

vacuum state, that is not symmetric in the positive and negative spectrum despite a so far symmetric theory.

However, there is an equivalent way to fix the negative spectrum, without changing the vacuum state. In the previous case, the operator $c^\dagger$ acts as annihilation operator and c as creation operator on the new vacuum:

$$c^\dagger|\text{vac}\rangle = c^\dagger(c^\dagger|0\rangle) = 0 \quad \text{and} \quad c|\text{vac}\rangle = cc^\dagger|0\rangle = |0\rangle \neq 0 \ .$$

So we define $c^\dagger = b$ and $c = b^\dagger$. Also, we had to exchange $c^\dagger c$ by $c^\dagger c - 1$ in the expression for $\widehat{\mathbf{H}}$. Hence:

$$\widehat{\mathbf{H}} = h(bb^\dagger - 1) = h((1 - b^\dagger b) - 1) = -hb^\dagger b \ .$$

By construction, $\widehat{\mathbf{H}}|0\rangle = 0$ and $\widehat{\mathbf{H}}b^\dagger|0\rangle = -h|0\rangle$, so $\widehat{\mathbf{H}} > 0$.

Definition 4.1.1.
The process of **normal ordering** $:\mathcal{O}:$ of an expression $\mathcal{O}$ is the reordering of operators, such that the creation operators are further left than the annihilation operators. For bosons a and fermions c normal ordering has the following rules:

i) $:a_i^\dagger a_j: = a_i^\dagger a_j$ and $:c_i^\dagger c_j: = c_i^\dagger c_j$.

ii) $:a_i a_j^\dagger: = a_j^\dagger a_i$ and $:c_i c_j^\dagger: = -c_j^\dagger c_i$.

iii) creation operators /annihilation operators change positions among themselves as usual (ccr- and car-relations).

iv) normal ordering is extended linearly.

With the concept of normal ordering, the construction of $\widehat{\mathbf{H}}$ can be realized as $:H:$. Choosing a polarization, such that creation and annihilation operators exchange their roles on the negative spectrum, i.e. H_-, then:

$$:H_-: = :hbb^\dagger: = -hb^\dagger b = \widehat{\mathbf{H}} \ .$$

4.2. Stable second quantization

Let $\mathcal{H}$ be a Hilbert space and $H \in \text{End}(\mathcal{H})$ with neither lower nor upper boundary for its spectrum and no zero eigenvalue, e.g. the Dirac operator. Also assume the existence of eigen states that define a Hilbert basis. Instead of choosing the natural polarization $\mathcal{H} \oplus \mathcal{H}^*$, we are looking for a subspace $P \subset \mathcal{H} \oplus \mathcal{H}^*$, such that the normal ordered second quantized operator $:\mathbf{H}:$ is non-negative with respect to that polarization.

By the choice of H, there are orthogonal projections Π_+ and Π_-, with

$$\mathcal{H} = \mathcal{H}_+ \oplus \mathcal{H}_- \ , \quad \text{with} \quad \mathcal{H}_\pm = \Pi_\pm \mathcal{H} \ .$$

Let $|h_j^+\rangle$ be the eigen states to positive eigen values and $|h_j^-\rangle$ those to negative eigen values, then $\Pi_+ \pm$ can be written explicitly:

$$\Pi_\pm = \sum_j |h_j^\pm\rangle\langle h_j^\pm| \ .$$

In other words, the Hilbert space has been decomposed into two orthogonal subspaces, one consisting of positive eigen states and their linear combinations, and one consisting of negative eigen states and their linear combinations.

Lemma 4.2.1.
Define $H_\pm := \Pi_\pm H \Pi_\pm$*, then it holds that:*

$$H = H_+ + H_- \ .$$

Proof 4.2.2.
Let $h_j^\pm$ be the eigen value to $|h_j^\pm\rangle$, then H can be written as follows:

$$H = \sum_j h_j^+ |h_j^+\rangle\langle h_j^+| \ + \ \sum_j h_j^- |h_j^-\rangle\langle h_j^-| \ .$$

Since $\langle h_j^+ | h_k^- \rangle = 0$ for all j, k we find for H_+:

$$\begin{aligned} H_+ = \Pi_+ H \Pi_+ &= \sum_{j,k,\ell} h_j^+ |h_k^+\rangle\langle h_k^+|h_j^+\rangle\langle h_j^+|h_\ell^+\rangle\langle h_\ell^+| \\ &= \sum_{j,k,\ell} h_j^+ \delta_{jk}\delta_{j\ell} |h_k^+\rangle\langle h_\ell^+| = \sum_j h_j^+ |h_j^+\rangle\langle h_j^+| \ . \end{aligned}$$

In the same way:

$$H_- = \sum_j h_j^- |h_j^-\rangle\langle h_j^-| \ ,$$

which proves the claim. □

Definition 4.2.3.
The **stable second quantization** is the second quantization with respect to the polarization

$$P = \mathcal{H}_+ \oplus \mathcal{H}_-^* \quad , \qquad P^* = \mathcal{H}_+^* \oplus \mathcal{H}_- \ .$$

The corresponding ferminoic Fock space is $\bigwedge(P)$.

Remembering remark 2.4.8, the second quantized operator $\mathbf{H}$ equals the representation of the corresponding Clifford algebra element H^C. For that reason, it is worthwhile to look at the representation again. Let $|e_i^\pm\rangle$ be Hilbert bases of $\mathcal{H}_\pm$ and $\langle e_i^\pm|$ the corresponding dual bases of $\mathcal{H}_\pm^*$. Further, let $|v^+\rangle \in \mathcal{H}_+$ and $\langle w^-| \in \mathcal{H}_-^*$, then the representation $\mathcal{D}_P$ works as follows:

$$\mathcal{D}_P(|e_j^+\rangle)\left(|v^+\rangle + \langle w^-|\right) = |e_j^+\rangle \wedge |v^+\rangle + |e_j^+\rangle \wedge \langle w^-| \ ,$$

$$\mathcal{D}_P(\langle e_k^+|)\left(|v^+\rangle + \langle w^-|\right) = Q\left(\langle e_k^+|, |v^+\rangle + \langle w^-|\right) = \langle e_k^+|v^+\rangle + 0 \ .$$

Similar rules apply to $\mathcal{D}_P(\langle e_k^-|)$ and $\mathcal{D}_P(|e_j^-\rangle)$.

$$\mathcal{D}_P(\langle e_j^-|)\left(|v^+\rangle + \langle w^-|\right) = \langle e_j^-| \wedge |v^+\rangle + \langle e_j^-| \wedge \langle w^-| \ ,$$
$$\mathcal{D}_P(|e_k^-\rangle)\left(|v^+\rangle + \langle w^-|\right) = Q\left(|e_k^-\rangle, |v^+\rangle + \langle w^-|\right) = 0 + \langle w^-|e_j^-\rangle \ .$$

It can be observed, that elements in $P = \mathcal{H}_+ \oplus \mathcal{H}_-^*$ still create and elements in $P^* = \mathcal{H}_+^* \oplus \mathcal{H}_-$ still annihilate. Also annihilation is again generalized as anti-derivative to $\bigwedge(P)$. To keep the notation short, we adopt the creation-/annihilation operator notation:

$$\mathcal{D}_P(|e_j^+\rangle) \equiv c_{+,j}^\dagger \ , \qquad \mathcal{D}_P(\langle e_j^+|) \equiv c_{+,j} \ , \qquad \mathcal{D}_P(\langle e_j^-|) \equiv c_{-,j}^\dagger \ , \qquad \mathcal{D}_P(|e_k^-\rangle) \equiv c_{-,j} \ .$$

Thus, the Clifford algebra form of H is:

$$\begin{aligned} H = H_+ + H_- &= \sum_{j,k} |e_j^+\rangle\langle e_j^+|H_+|e_k^+\rangle\langle e_k^+| + \sum_{j,k} |e_j^-\rangle\langle e_j^-|H_-|e_k^-\rangle\langle e_k^-| \\ &\equiv \sum_{j,k} (H_+)_{jk}|e_j^+\rangle\langle e_k^+| + \sum_{j,k} (H_-)_{jk}|e_j^-\rangle\langle e_k^-| \end{aligned}$$

$$\Rightarrow \qquad H^C = \sum_{j,k} (H_+)_{jk}|e_j^+\rangle \circ \langle e_k^+| + \sum_{j,k} (H_-)_{jk}|e_j^-\rangle \circ \langle e_k^-|$$

$$\Rightarrow \qquad \mathcal{D}_P(H^C) = \sum_{j,k} (H_+)_{jk}\ c_{+,j}^\dagger c_{+,k} \ + \ \sum_{j,k} (H_-)_{jk}\ c_{-,j} c_{-,k}^\dagger$$

The normal ordered second quantized operator is:

$$\boxed{:\mathbf{H}: = :\mathcal{D}_P\left(H^C\right): = \sum_{j,k} (H_+)_{jk}\ c_{+,j}^\dagger c_{+,k} \ - \ \sum_{j,k} (H_-)_{jk}\ c_{-,k}^\dagger c_{-,j} \ .}$$

Remark 4.2.4.
Since normal ordering has the prerequisite of second quantization (here), we will drop the special notation for second quantization and only write $:H:$

Lemma 4.2.5.
The operator $:H:$ is a derivation on $\bigwedge(P)$.

Proof 4.2.6.
The first step is to observe, that the pair of creation and annihilation operators $(c_{\pm,j}^\dagger c_{\pm,k})$ is a derivation on $\bigwedge(P)$. To see that, let $v \in P$ and $\xi \in \bigwedge(P)$, then:

$$\begin{aligned} c_{+,j}^\dagger c_{+,k}(v \wedge \xi) &= c_{+,j}^\dagger \left((c_{+,k}v) \wedge \xi - v \wedge c_{+,k}\xi\right) \\ &= (c_{+,j}^\dagger c_{+,k}v) \wedge \xi c_{+,j}^\dagger (v \wedge c_{+,k}\xi) \\ &= (c_{+,j}^\dagger c_{+,k}v) \wedge \xi + v \wedge (c_{+,j}^\dagger c_{+,k}\xi) \ . \end{aligned}$$

The same holds true for $c_{-,j}^\dagger c_{-,k}$, which proves the claim by linearity. □

Theorem 4.2.7.
The operator $:H:$ *is a positive on* $\bigwedge(P)$.

Proof 4.2.8.
The operator $:H:$ is a derivation on $\bigwedge(P)$. Hence it is enough to show positivity on $P = \mathcal{H}_+ \oplus \mathcal{H}_-^*$. On $\mathcal{H}_+$ the operator $:H:$ acts as H_+, which is positive by definition and on $\mathcal{H}_-^*$ it acts as $-H_-^t$, being also positive, because of the minus sign. □

4.3. Stable second quantization of the Dirac Hamiltonian

To apply the stable second quantization to the free Dirac Hamiltonian, which does not have a discrete spectrum, we understand the meaning of eigen vectore more liberally, and ignore mathematical subtleties. Also, for simplicity we follow [Zir10], that is, we fix an inertial frame and set $t = 0$, not using the relativistic invariance.

In this case, the single particle Hilbert space $\mathcal{H}$ is the space of spinor fields. Since we will chosen a basis anyway, this space is (not canonnically) isomorphic to $\mathcal{H} \cong L^2(\mathbb{R}^3) \otimes \mathbb{C}^4$. Thus, H becomes a matrix $(H_{\tau\epsilon})$. In the momentum space, for fixed k, the Dirac equation does not contain any derivatives and becomes a usual matrix equation.

$$H(k)\psi(k) = \begin{pmatrix} mc^2 & \hbar c \sum_j \sigma_j k_j \\ \hbar c \sum_j \sigma_j k_j & -mc^2 \end{pmatrix} \psi(k) = E(k)\psi(k) \ .$$

The square of $H(k)$ is $(\hbar\omega(k))^2 \mathbb{1}$, where

$$\hbar\omega(k) = \sqrt{m^2c^4 + \hbar^2 k^2 c^2} \ ,$$

i.e. the energy momentum relation. Then it follows that:

$$H(k)^2\psi(k) = (\hbar\omega(k))^2\psi(k) = E(k)\ H(k)\psi(k) = E(k)^2\psi(k) \ .$$

This means, the eigen values are $\pm\hbar\omega(k)$. For the two eigen values, we define eigen states:

$$H(k)u_s(k) = +\hbar\omega(k)\ u_s(k) \qquad \text{and} \qquad H(k)v_s(k) = -\hbar\omega(k)\ v_s(k) \ ,$$

where the index $s \in \{+, -\}$ denotes the two-fold spin degeneracy for each eigen value. Here we follow [Zir14] and choose the non-conventional but simpler normalization conditions:

$$u_s^\dagger(k)u_{s'}(k) = \delta_{ss'} = v_s^\dagger(k)v_{s'}(k) \qquad \text{and} \qquad u_s^\dagger(k)v_{s'}(k) = 0 \ .$$

The hermitian adjoint is to be understood as the adjoint with respect to the hermitian structure of $\mathbb{C}^4$, defining a bijection to the dual space. For the spin degeneracy a common choice is, that s denotes the helecity, i.e. the eigen value with respect to the operator (see lemma 3.4.1) $\frac{1}{|k|}\sum_\ell \mathcal{D}_S(\Sigma_\ell)k_\ell$.

By the choice of the eigen states, and with the isomorphism between linear maps and tensors, we can write $\Pi_\pm(k)$ explicitly:

$$\Pi_+(k) = \sum_s u_s(k) \otimes u_s^\dagger(k) \qquad \text{and} \qquad \Pi_-(k) = \sum_s v_s(k) \otimes v_s^\dagger(k) \ ,$$

$$\Rightarrow \qquad \mathbb{1} = \sum_s u_s(k) \otimes u_s^\dagger(k) + \sum_s v_s(k) \otimes v_s^\dagger(k) \ . \tag{4.1}$$

The procedure of stable second quantization now proceeds by turning these operators into the corresponding Clifford algebra elements and to represent them as operators on the exterior algebra. To be more transparent, we reiterate the steps:

1. First, we define the positive /negative subspaces $\mathcal{H}_\pm = \Pi_\pm \mathcal{H}$, and choose the Fockspace $\bigwedge(\mathcal{H}_+ \oplus \mathcal{H}_-^*)$.

2. The choice of eigen states for $H(k)$ defines a basis for fixed k such that we define for the positive subspace:

   $$c_{+,s}^\dagger(k) \equiv \mathcal{D}_P(u_s(k)) \qquad \text{and} \qquad c_{+,s}(k) \equiv \mathcal{D}_P(u_s^\dagger(k)) \ .$$

   Here, there is no mistake in the positions of $\dagger$, since $u^\dagger$ is an element of the dual space and thus corresponds to annihilation.

3. In the stable second quantization, the Hamilton operator acts as $-H^t$ on the negative subspace, which fixes the negative spectrum. However, then $\mathcal{D}_P(v_s^\dagger(k))$ creates a particle with negative momentum $-k$ since $p \to -p^t$. Furthermore, the created particle has also the helecity $-s$ for the same reason. To keep the notation meaningful, we have to define:

   $$c_{-,s}^\dagger(k) \equiv \mathcal{D}_P(v_{-s}^\dagger(-k)) \qquad \text{and} \qquad c_{-,s}(k) \equiv \mathcal{D}_P(v_{-s}(-k)) \ .$$

4. The Hamilton operator for fixed k is:

   $$:H(k): = \hbar\omega(k) \cdot \Pi_+(k) + (-\hbar\omega(k)) \cdot \Pi_-(k) = \hbar\omega(k)(\Pi_+(k) - \Pi_-(k) \ .$$

   Applying the the stable second quantization and summing/integrating over the momentum space with measure $\frac{dk^3}{(2\pi)^3}$ yields:

   $$\begin{aligned} :H: &= \int \frac{dk^3}{(2\pi)^3} \ \hbar\omega(k) \ \sum_s \left(c_{+,s}^\dagger(k) c_{+,s}(k) + c_{-,-s}^\dagger(-k) c_{-,-s}(-k) \right) \\ &= \int \frac{dk^3}{(2\pi)^3} \ \hbar\omega(k) \ \sum_s c_{+,s}^\dagger(k) c_{+,s}(k) \\ &\qquad + \int \frac{dk^3}{(2\pi)^3} \ \hbar\omega(k) \ \sum_s c_{-,-s}^\dagger(-k) c_{-,-s}(-k) \\ &= \int \frac{dk^3}{(2\pi)^3} \ \hbar\omega(k) \ \sum_s \left(c_{+,s}^\dagger(k) c_{+,s}(k) + c_{-,s}^\dagger(k) c_{-,s}(k) \right) \ . \end{aligned}$$

   In the last equality we used, that we sum over all s and integrate over all k, where $\omega(k) = \omega(-k)$ allowing to ignore the minus sign. [1]

[1] Œ in one dimension:

$$\int_{-\infty}^{\infty} dk \ \omega(k) f(-k) = \int_{\infty}^{-\infty} d(-k) \ \omega(-k) f(k) = -\int_{-\infty}^{\infty} d(-k) \ \omega(-k) f(k)$$

Lemma 4.3.1.
The ***electron/positron creation /annihilation operators*** $c^\dagger_{\pm,s}(k)$ *and* $c_{\pm,s}(k)$ *satisfy the following anti-commutator relations:*

$$\{c_{\alpha,s}(k), c^\dagger_{\beta,s'}(k')\} = (2\pi)^3 \delta_{\alpha\beta}\delta_{ss'}\delta(k-k') \ ,$$

$$\{c^\dagger_{\alpha,s}(k), c^\dagger_{\beta,s'}(k')\} = 0 = \{c_{\alpha,s}(k), c_{\beta,s'}(k')\} \ .$$

Proof 4.3.2.
The operators $c^\dagger_{\pm,s}(k)$ and $c_{\pm,s}(k)$ have a momentum dependence. Let $c^\dagger_{\pm,s}(x)$ and $c_{\pm,s}(x)$ be the corresponding position dependent operators, related by Fourier transformation:

$$c_{\pm,s}(k) = \int dx^3 \ e^{-ikx} c_{\pm,s}(x) \qquad \text{and} \qquad c^\dagger_{\pm,s}(k) = \int dx^3 \ e^{ikx} c^\dagger_{\pm,s}(x) \ .$$

The operator fields have to satisfy the canonical field anti-commutator relations (as used in canonical quantizations):

$$\{c_{\alpha,s}(x), c^\dagger_{\beta,s'}(x')\} = \delta_{ss'}\delta_{\alpha\beta}\delta(x-x') \ ,$$

$$\{c^\dagger_{\alpha,s}(x), c^\dagger_{\beta,s'}(x')\} = 0 = \{c_{\alpha,s}(x), c_{\beta,s'}(x')\} \ .$$

The Kronecker-deltas $\delta_{ss'}$ and $\delta_{\alpha\beta}$ are motivated by theorem D.4.4 and the fact, that different indices correspond to different "elements" in the Hilbert space.

Hence, using lemma F.3.4 in the integral kernel notation together with quantum mechanical Fourier normalization:

$$\begin{aligned}
\{c_{\pm,s}(k), c^\dagger_{\pm,s'}(k')\} &= \int dx^3 \int dy^3 \ e^{-ikx} e^{ik'y} \{c_{\pm,s}(x), c^\dagger_{\pm,s'}(y)\} \\
&= \delta_{ss'} \int dx^3 \int dy^3 \ e^{-ikx} e^{ik'y} \delta(x-y) \\
&= \delta_{ss'} \int dx^3 \ e^{ix(k'-k)} = (2\pi)^3 \delta_{ss'}\delta(k-k') \ .
\end{aligned}$$

The remaining anti-commutator relations can be shown in the same way. □

4.4. Dirac field

So far, we have only found an expression for the Hamiltonian, that fixes the issues of the one particle Dirac theory. One could try to apply the stable second quantization for other operators, trying to find mutual compatible splittings $\mathcal{H} = \mathcal{H}_+ \oplus \mathcal{H}_-$. Yet, since a modern description is a field theoretic description anyway, we will adopt a concept

$$= -\int_{-\infty}^{\infty} -dk \ \omega(k) f(k) = \int_{-\infty}^{\infty} dk \ \omega(k) f(k) \ .$$

called field operator. Here we introduce it from the perspective of second quantization, rather then canonical quantization (since it yields the same result in this case).

Let $|j\rangle$ be a proper Hilbert basis, $|x\rangle$ and $|y\rangle$ be position states respectively, then, using the pseudo completeness relations:

$$H = \sum_{k,j} |j\rangle\langle j|H|k\rangle\langle k| \qquad \rightsquigarrow \qquad \mathbf{H} = \sum_{j,k} \langle j|H|k\rangle c_j^\dagger c_k$$

$$\begin{aligned}\mathbf{H} &= \sum_{j,k} \langle j|H|k\rangle c_j^\dagger c_k = \int dx^3 \int dy^3 \sum_{j,k} \langle j|y\rangle\langle y|H|x\rangle\langle x|k\rangle c_j^\dagger c_k \\ &= \int dx^3 \int dy^3 \ \langle y|H|x\rangle \left(\sum_j c_j^\dagger \langle j|y\rangle\right)\left(\sum_k \langle x|k\rangle c_k\right) \\ &\equiv \int dx^3 \int dy^3 \ \langle y|H|x\rangle \Psi^\dagger(y)\Psi(x) \ .\end{aligned}$$

Definition 4.4.1.
The **field operators** are defined by

$$\Psi^\dagger(x) = \sum_j c_j^\dagger \langle j|x\rangle \qquad \text{and} \qquad \Psi(x) = \sum_j \langle x|j\rangle c_j \ ,$$

where $\{|j\rangle\}$ is a Hilbert basis of the single particle Hilbert space and $c_j^\dagger$, c_j are the corresponding bosonic/fermionic creation and annihilation operators.

With definition E.2.2, the formula for the second quantized operator can be simplified:

$$\begin{aligned}\mathbf{H} &= \int dx^3 \int dy^3 \ \Psi^\dagger(y)\langle y|H|x\rangle\Psi(x) = \int dx^3 \int dy^3 \ \Psi^\dagger(y)\langle y|x\rangle H\Psi(x) \\ &= \int dx^3 \int dy^3 \ \Psi^\dagger(y)\delta(x-y)H\Psi(x) = \int dx^3 \Psi^\dagger(x)H\Psi(x) \ ,\end{aligned}$$

where H acts on $\Psi(x)$ by its representation on $L^2(\mathbb{R}^n)$.

Lemma 4.4.2.
The Field operators satisfy the following (anti)-commutation relations:

$$[\Psi^\dagger(x), \Psi^\dagger(y)]_\pm = 0 \ , \qquad [\Psi(x), \Psi(y)]_\pm = 0$$

$$\textit{and} \quad [\Psi(x), \Psi^\dagger(y)]_\pm = \delta(x-y) \ .$$

Proof 4.4.3.
This statement seems simple enough. Yet, to exchange the order of sums, it is necessary to show, that the limits behind the infinite sums are continuous with respect to a norm/topology, we are yet to choose. Here we do assume this to be true.

$$[\Psi(x), \Psi(y)]_\pm = \sum_{j,k} \langle x|j\rangle\langle y|k\rangle [c_j, c_k]_\pm = \sum_{j,k} \langle x|j\rangle\langle y|k\rangle \cdot 0 = 0 \ .$$

The same calculation holds for $[\Psi^\dagger(x), \Psi^\dagger(y)]_\pm$. For the last relation, we are sloppy again, using $\langle x|y\rangle = \delta(x-y)$:

$$\begin{aligned}[\Psi(x), \Psi^\dagger(y)]_\pm &= \sum_{j,k} \langle x|j\rangle\langle k|y\rangle [c_j, c_k^\dagger]_\pm = \sum_{j,k} \langle x|j\rangle\langle k|y\rangle \delta_{jk} \\ &= \sum_j \langle x|j\rangle\langle j|y\rangle = \langle x|y\rangle = \delta(x-y) \ .\end{aligned}$$

□

Since we have already used the pseudo-completeness relations, we may as well drop the condition of proper Hilbert basis and define

$$\Psi(x) = \int \frac{dk^3}{(2\pi)^3} \langle x|k\rangle c(k) \ .$$

With (anti)-commutation relations for the creation/annihilation operators using Delta functions instead of Kronecker deltas, as in lemma 4.3.1. The properties for the Field operators presented so far stay the same, as can be checked by similar calculations, using the pseudo completeness relations instead of the proper completeness relations.

For the Dirac field we make the ansatz by a mode expansion of positive/negative energy states $\Psi(x) = \Psi^+(x) + \Psi^-(x)$:

$$\Psi^+(x) = \int \frac{dk^3}{(2\pi)^3} e^{ikx} \sum_s u_s(k) c_{+,s}(k)$$

$$\Psi^-(x) = \int \frac{dk^3}{(2\pi)^3} e^{ikx} \sum_s v_s(k) c^\dagger_{-,-s}(-k) \ .$$

The peculiar choice of $c^\dagger_{-,-s}(k)$ is because we already know the meaning of creation/annihilation operators from the stable second quantization ($c_{+,s}(k) \equiv \mathcal{D}_P(u_s^\dagger(k))$ and $c^\dagger_{-,s}(k) \equiv \mathcal{D}_P(v^\dagger_{-s}(-k))$).

Definition 4.4.4.
The **Dirac field** is defined by

$$\Psi(x) = \int \frac{dk^3}{(2\pi)^3} e^{ikx} \sum_s \left(u_s(k) c_{+,s}(k) + v_s(k) c^\dagger_{-,-s}(-k) \right) \ .$$

Choosing different spinors $u_s(k)$ and $v_s(k)$, the Dirac field might need different normalization factors, such that one obtains the anti commutation relations form the next lemma.

Remark 4.4.5.
In quantum field theory, using the canonical quantizations explained in chapter 5, one obtains the same field operators (after applying Dirac's hole theory). Using Noether's theorem, one obtains the same expressions for the momentum and charge operators. In this case, the energy-momentum tensor and $U(1)$-phase symmetry for charge are used.

For completeness, it should be mentioned, that in the process of canonical quantization one obtains different creation/annihilation operators, whose physical meaning has to be deduced.

Lemma 4.4.6.
The Dirac field satisfies the canonical field anti commutator relations:

$$\{\Psi^\dagger{}_\alpha(x), \Psi^\dagger{}_\beta(y)\} = 0 \ , \qquad \{\Psi^\alpha(x), \Psi^\beta(y)\} = 0$$

$$\textit{and} \quad \{\Psi^\alpha(x), \Psi^\dagger{}_\beta(y)\} = \delta^\alpha_\beta \delta(x-y) \ .$$

Proof 4.4.7.
First calculate $\Psi^\dagger(x)$:

$$\Psi^\dagger(x) = \int \frac{dk^3}{(2\pi)^3} e^{-ikx} \sum_s \left(c^\dagger_{+,s}(k) u^\dagger_s(k) + c_{-,-s}(-k) v^\dagger_s(k) \right) \ .$$

To avoid an overload of notation we write $u_\alpha(s,k)$ instead of $(u_s(k))_\alpha$. Then, with lemma 4.3.1 we find:

$$\begin{aligned}
&\Big\{ \left(u^\alpha(s,k) c_{+,s}(k) + v^\alpha(s,k) c^\dagger_{-,-s}(-k) \right), \\
&\qquad \left(c^\dagger_{+,r}(\ell) u^\dagger{}_\beta(r,\ell) + e^{i\ell y} c_{-,-r}(-\ell) v_\dagger{}^\beta(r,\ell) \right) \Big\} \\
&= \left\{ u^\alpha(s,k) c_{+,s}(k), c^\dagger_{+,r}(\ell) u^\dagger{}_\beta(r,\ell) \right\} \\
&\qquad + \left\{ v^\alpha(s,k) c^\dagger_{-,-s}(-k), c_{-,-r}(-\ell) v^\dagger{}_\beta(r,\ell) \right\} + 0 \\
&= u^\alpha(s,k) u^\dagger{}_\beta(r,\ell) \left\{ c_{+,s}(k), c^\dagger_{+,r}(\ell) \right\} \\
&\qquad + v^\alpha(s,k) v^\dagger{}_\beta(r,\ell) \left\{ c^\dagger_{-,-s}(-k), c_{-,-r}(-\ell) \right\} \\
&= \left(u^\alpha(s,k) u^\dagger{}_\beta(r,\ell) + v^\alpha(s,k) v^\dagger{}_\beta(r,\ell) \right) \delta_{rs} \delta(k-\ell)
\end{aligned}$$

Finally plugging in and using (4.1), which reads
$\sum_s u_\alpha(s,k) u^\dagger{}_\beta(s,k) + v^\alpha(s,k) v^{\dagger\beta}(s,k) = \delta^\alpha_\beta$:

$$\begin{aligned}
\{\Psi^\alpha(x), \Psi^\dagger{}_\beta(y)\} &= \int \frac{dk^3}{(2\pi)^3} \int \frac{d\ell^3}{(2\pi)^3} \, e^{ikx} e^{-i\ell y} \sum_{r,s} \Big\{ \left(u^\alpha(s,k) c_{+,s}(k) + v^\alpha(s,k) c^\dagger_{-,-s}(-k) \right), \\
&\qquad \left(c^\dagger_{+,r}(\ell) u^\dagger{}_\beta(r,\ell) + c_{-,-r}(-\ell) v^\dagger{}_\beta(r,\ell) \right) \Big\} \\
&= \int \frac{dk^3}{(2\pi)^3} \int \frac{d\ell^3}{(2\pi)^3} \, e^{ikx} e^{-i\ell y} \sum_{r,s} \Big(u^\alpha(s,k) u^\dagger{}_\beta(r,\ell) \\
&\qquad + v^\alpha(s,k) v^\dagger{}_\beta(r,\ell) \Big) \delta_{rs} \delta(k-\ell) \\
&= \int \frac{dk^3}{(2\pi)^3} e^{ikx} e^{-iky} \sum_s u^\alpha(s,k) u^\dagger{}_\beta(s,k) + v^\alpha(s,k) v^\dagger{}_\beta(s,k) \\
&= \delta^\alpha_\beta \int \frac{dk^3}{(2\pi)^3} e^{ik(x-y)} = \delta^\alpha_\beta \delta(x-y) \ .
\end{aligned}$$

The remaining anti-commutator relations can be shown in the same way. □

Theorem 4.4.8.
Let H be the Dirac Hamiltonian, then the stable second quantized Hamiltonian $:H:$ is given by:

$$:H: = :\int dx^3 \Psi^\dagger(x) H \Psi(x): \ .$$

Proof 4.4.9.
First we need to evaluate $H\Psi(x)$. We notice, that H in the position representation is given by

$$\begin{pmatrix} mc^2 & -i\hbar c \sum_j \sigma_j \partial_{x_j} \\ -i\hbar c \sum_j \sigma_j \partial_{x_j} & -mc^2 \end{pmatrix} ,$$

if we use the $\mathbb{C}^4$ components. This means:

$$\begin{aligned} H \ e^{ikx} \sum_s \Big(u_s(k) c_{+,s}(k) + v_s(k) c^\dagger_{-,-s}(-k) \Big) \\ = e^{ikx} \sum_s H(k) \Big(u_s(k) c_{+,s}(k) + v_s(k) c^\dagger_{-,-s}(-k) \Big) \ , \end{aligned}$$

where $H(k)$ is the matrix of H in the momentum pace. Thus

$$H\Psi(x) = \int \frac{dk^3}{(2\pi)^3} e^{ikx} \hbar\omega(k) \sum_s \Big(u_s(k) c_{+,s}(k) - v_s(k) c^\dagger_{-,-s}(-k) \Big) \ .$$

For $\int dx^3 \Psi^\dagger(x) H \Psi(x)$ we find:

$$\begin{aligned} \int dx^3 \Psi^\dagger(x) H \Psi(x) &= \int dx^3 \int \frac{d\ell^3}{(2\pi)^3} \int \frac{dk^3}{(2\pi)^3} e^{-i\ell x} e^{ikx} \hbar\omega(k) \\ &\quad \sum_{r,s} \Big(c^\dagger_{+,r}(\ell) u^\dagger_r(\ell) u_s(k) c_{+,s}(k) - c_{-,-r}(-\ell) v^\dagger_r(\ell) v_s(k) c^\dagger_{-,-s}(-k) \\ &\quad - c^\dagger_{+,r}(\ell) u^\dagger_r(\ell) v_s(k) c^\dagger_{-,-s}(-k) + c_{-,-r}(-\ell) v^\dagger_r(\ell) u_s(k) c_{+,s}(k) \Big) \\ &= \int \frac{d\ell^3}{(2\pi)^3} \int \frac{dk^3}{(2\pi)^3} \ \Big(\int dx^3 \ e^{i(k-\ell)x} \Big) \hbar\omega(k) \\ &\quad \sum_{r,s} \Big(c^\dagger_{+,r}(\ell) c_{+,s}(k) - c_{-,-r}(-\ell) c^\dagger_{-,-s}(-k) \Big) \delta_{rs} \\ &= \int \frac{d\ell^3}{(2\pi)^3} \int \frac{dk^3}{(2\pi)^3} \ (2\pi)^3 \delta(k-\ell) \hbar\omega(k) \\ &\quad \sum_{s} \Big(c^\dagger_{+,s}(\ell) c_{+,s}(k) - c_{-,-s}(-\ell) c^\dagger_{-,-s}(-k) \Big) \\ &= \int \frac{dk^3}{(2\pi)^3} \ \hbar\omega(k) \sum_s \Big(c^\dagger_{+,s}(\ell) c_{+,s}(k) - c_{-,s}(k) c^\dagger_{-,s}(k) \Big) \ . \end{aligned} \tag{4.2}$$

In the last step, we have changed the sign of $-k$ and $-s$, for the same reason as in the last subsection. Normal ordering concludes the proof:

$$:\int dx^3 \Psi^\dagger(x) H \Psi(x): = \int \frac{dk^3}{(2\pi)^3}\ \hbar\omega(k)\ \sum_s \left(c^\dagger_{+,s}(k)c_{+,s}(k) + c^\dagger_{-,s}(k)c_{-,s}(k)\right) = :H:\ .$$

□

With similar calculations, the momentum and charge can be calculated:

$$:p_j: = :\int dx\ \Psi^\dagger(x) p_j \Psi(x): = \int \frac{dk^3}{(2\pi)^3}\ \hbar k_j\ \sum_s \left(c^\dagger_{+,s}(k)c_{+,s}(k) + c^\dagger_{-,s}(k)c_{-,s}(k)\right)$$

$$:Q: = :e\cdot\int dx\ \Psi^\dagger(x)\Psi(x): = e\cdot\int \frac{dk^3}{(2\pi)^3}\ \sum_s \left(c^\dagger_{+,s}(k)c_{+,s}(k) - c^\dagger_{-,s}(k)c_{-,s}(k)\right)$$

Remark 4.4.10.
The reason for $:p_j:$ to have the same sign for electrons c_+ and positrons c_- is, that k_j unlike $\omega(k)$ is not symmetric but anti symmetric. So to go from $c_{-,s}(-k)$ to $c_{-,s}(k)$ etc. the sign in front changes. For $:Q:$ however, there is nothing of that sort, such that electrons and positrons have different charge. This is the reinterpretation of the probability density as charge density.

4.5. Dirac field and time dependence

One way to find the time dependence, one could quantize the covariant Dirac action. Here we choose the more intuitive way, by splitting the Dirac field in positive and negative part and adding the corresponding time exponential $e^{-\frac{i}{\hbar}E_k t}$, where $E_k = \pm\hbar\omega(k)$, from the plane wave solutions:

$$\begin{aligned}\Psi(x) &= \int \frac{dk^3}{(2\pi)^3} e^{ikx} \sum_s \left(u_s(k)c_{+,s}(k) + v_s(k)c^\dagger_{-,-s}(-k)\right)\\ &= \int \frac{dk^3}{(2\pi)^3} \sum_s \left(e^{ikx}u_s(k)c_{+,s}(k) + e^{-ikx}v_s(-k)c^\dagger_{-,-s}(k)\right)\end{aligned}$$

$$\rightsquigarrow \quad \Psi(x,t) = \int \frac{dk^3}{(2\pi)^3} \sum_s \left(e^{ikx-i\omega(k)t}u_s(k)c_{+,s}(k) + e^{-ikx+i\omega(k)t}v_s(-k)c^\dagger_{-,-s}(k)\right)\ .$$

Next, we use the four-wave-vector $\mathbf{k} = (\frac{\omega(k)}{c}, k)$ and the four position $\mathbf{x} = (ct, x)$, together with the Minkowski metric $\mathbf{kx} = g(\mathbf{k},\mathbf{x}) = \omega(k)t - kx$ to find:

$$\Psi(\mathbf{x}) = \int \frac{dk^3}{(2\pi)^3} \sum_s \left(e^{-i\mathbf{kx}}u_s(k)c_{+,s}(k) + e^{i\mathbf{kx}}v_s(-k)c^\dagger_{-,-s}(k)\right)\ .$$

4.6. Relativistic conventions and useful relations

In the last sections, the relativistic covariance of the Dirac theory has not been emphasized, to keep the focus on the many particle interpretation. However, as most textbooks use the relativistic formulation with appropriate conventions (compare for example [BD94, section 3.1]), it is worthwhile to consider them as well.

To adopt the conventions of [BD94, section 3.1] we follow the remarks of [Zir14, appendix of the quantization of the Dirac field].

4.6.1. Useful relations of the γ-matrices

Recalling subsection 3.3.2, the hermiticity of β and α_ℓ lead to a continuity equation, that corresponds to charge conservation in the many particle interpretation of the Dirac theory. The hermiticity so far is a result of the Dirac representation of the γ-matrices and (3.2) This hermiticity is promoted to an axiom of the theory. For the γ-matrices, this axiom reads:

$$\boxed{(\gamma^\mu)^\dagger = \gamma^0\gamma^\mu\gamma^0 \ .}$$

Using (3.1), this can be seen as follows:

$$\beta = \beta^\dagger = \left(\gamma^0\right)^\dagger = \gamma^0\gamma^0\gamma^0 = \beta\beta^2 = \beta \ ,$$

$$\left(\gamma^j\right)^\dagger = (\beta\alpha_j)^\dagger = \alpha_j^\dagger\beta^\dagger = \alpha_j\beta = \beta^2\alpha_j\beta = \beta(\beta\alpha_j)\beta = \gamma^0\gamma^j\gamma^0 \ .$$

For the trace identities of gamma matrices, that do not depend on the chosen representation, we introduce a shorthand notation, often used in field theory and relativistic quantum mechanics:

Definition 4.6.1 (Feynman slash notation)**.**
Let $A = (A_0, A_1, A_2, A_3)$ be a family of four objects, then

$$\not{A} := A_\mu\gamma^\mu \ .$$

Another convention is the definition of γ^5:

$$\boxed{\gamma^5 := i\gamma^0\gamma^1\gamma^2\gamma^3 \ .}$$

Remark 4.6.2.
We make the convention here, that with gamma matrices we only mean the original four γ^μ.

Corollary 4.6.3.
γ^5 has the following properties:

$$i) \quad \left(\gamma^5\right)^\dagger = \gamma^5 \qquad ii) \quad \left(\gamma^5\right)^2 = \mathbb{1} \qquad iii) \quad \{\gamma^5, \gamma^\mu\} = 0 \ .$$

Proof 4.6.4.

i) Using $(\gamma^0)^2 = \mathbb{1}$:

$$\begin{aligned}\left(\gamma^5\right)^\dagger &= -i(\gamma^0)^\dagger(\gamma^1)^\dagger(\gamma^1)^\dagger(\gamma^3)^\dagger = -i\gamma^0\gamma^0\gamma^1\gamma^0\gamma^0\gamma^2\gamma^0\gamma^0\gamma^1\gamma^0 \\ &= -i\gamma^1\gamma^2\gamma^3\gamma^0 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \gamma^5 \ .\end{aligned}$$

ii) Using that the gamma matrices anticommute and $(\gamma^0)^2 = \mathbb{1}$, $(\gamma^j)^2 = -\mathbb{1}$:

$$\left(\gamma^5\right)^2 = -\gamma^0\gamma^1\gamma^2\gamma^3\gamma^0\gamma^1\gamma^2\gamma^3 = -(-1)^6(-1)^3\mathbb{1} = \mathbb{1} \ .$$

iii) Since $\gamma^\mu\gamma^\nu = -\gamma^\nu\gamma^\mu$ for $\mu \neq \nu$ there are always 3 interchanges, such that

$$\gamma^5\gamma^\mu = (-1)^3\gamma^\mu\gamma^5 = -\gamma^\mu\gamma^5 \ .$$

□

We conclude this subsection by stating some representation independent identities for the γ-matrices. Most of the proofs can be found in [BD94, section 7.2].

- $\gamma^\mu\gamma^\mu = g_{\mu\mu}, \quad \gamma_\mu\gamma^\mu = 4\mathbb{1}, \quad \gamma_\mu \not{a}\gamma^\mu = -2\not{a}, \quad \gamma_\mu \not{a}\not{b}\gamma^\mu = 4a\cdot b\ \mathbb{1},$
  $\gamma_\mu \not{a}\not{b}\not{c}\gamma^\mu = -2\not{c}\not{b}\not{a}$ and $\gamma_\mu \not{a}\not{b}\not{c}\not{d}\gamma^\mu = 2(\not{d}\not{a}\not{b}\not{c} + \not{c}\not{b}\not{a}\not{d})$.
- $\mathrm{tr}(\gamma^{\mu_1}\dots\gamma^{\mu_n}) = 0$, $\mathrm{tr}(\gamma^5) = 0$ and $\mathrm{tr}(\gamma^5\gamma^{\mu_1}\dots\gamma^{\mu_n}) = 0$ for odd n.
- $\mathrm{tr}(\not{a}\not{b}) = 4a\cdot b\ \mathbb{1}$.
- $\mathrm{tr}(\not{a}_1\dots\not{a}_n) = \sum_{j=2}^{n}(-1)^j a_1\cdot a_j\ \mathrm{tr}(\not{a}_s\dots\widehat{\not{a}_j}\dots\not{a}_n)$.
- $\mathrm{tr}(\gamma^5\not{a}\not{b}) = 0$ and $\mathrm{tr}(\gamma^5\not{a}\not{b}\not{c}\not{d}) = 4i\varepsilon_{\alpha\beta\gamma\delta}a^\alpha b^\beta c^\gamma d^\delta\ \mathbb{1}$.
- $\mathrm{tr}(\not{a}_1\not{a}_2\dots\not{a}_{2n}) = \mathrm{tr}(\not{a}_{2n}\dots\not{a}_2\not{a}_1)$.

4.6.2. Construction of $u_s(k)$ and $v_s(k)$

To construct the Dirac field, positive and negative solutions with eigen spinors u_s and v_s have been used. The k dependence has not been specified up to the convenient normalization conditions $u_s^\dagger(k)u_{s'}(k) = \delta_{ss'} = v_s^\dagger(k)v_{s'}(k)$ and $u_s^\dagger(k)v_{s'}(k) = 0$. In the relativistic version, one constructs the k dependence from the static solution $k = 0$ by lorentz boosts, which will give rise to new normalization conditions. We start form the static situation $k = 0$ for the relativistic Dirac equation:

$$\left(\gamma^0\partial_t + \frac{imc^2}{\hbar}\right)\psi(x) = \begin{pmatrix} \partial_t + \frac{imc2}{\hbar} & 0 & 0 & 0 \\ 0 & \partial_t + \frac{imc^2}{\hbar} & 0 & 0 \\ 0 & 0 & -\partial_t + \frac{imc^2}{\hbar} & 0 \\ 0 & 0 & 0 & -\partial_t + \frac{imc^2}{\hbar} \end{pmatrix}\psi(x) = 0$$

As ususal, the ansatz is $\psi_r^{(0)}(\mathbf{x}) = w_r(0)e^{-\varepsilon_r \frac{imc^2}{\hbar}t} = w_r(0)e^{-\varepsilon_r \frac{imc}{\hbar}x^0}$ with spinors $w_r(0)$, where we have defined:

$$\varepsilon_r = \begin{cases} 1 & , \ r = 1,2 \quad \text{positive energy solutions} \\ -1 & , \ r = 3,4 \quad \text{negative energy solutions} \end{cases} .$$

For the positive/negative energy solutions, we choose the spinors (in Dirac representation):

$$w_1(0) = u_+(0) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} , \qquad w_2(0) = u_-(0) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} ,$$

$$w_3(0) = v_+(0) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} , \qquad w_4(0) = v_-(0) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} .$$

In general, not using the Dirac representation, one demands that $w_r(0)^\dagger \gamma^0 w_{r'}(0) = \varepsilon_r \delta_{rr'}$.

To define $\omega_r(k)$ for non-zero k, we consider a general solution $\psi_r(\mathbf{q}) = w_r(k)e^{-\varepsilon_r ik_\mu x^\mu(\mathbf{q})}$. The idea is now, to Lorentz boost the static solution $\psi_r^{(0)}(\mathbf{q})$ to a solution $\psi_r(\mathbf{q})$. Formally we consider a Lorentz boost (special Lorentz transformation) $L = e^X \in \mathrm{SO}(\mathbb{R}^4, g)$, with $X \in \mathfrak{so}(\mathbb{R}^4, g)$. From subsection D.5.2 and definition 3.2.11 we know, how Lorentz transformations act on spinor fields:

$$\mathcal{D}(L)\psi(\mathbf{q}) = \mathcal{D}_S\left(e^{\tau^{-1}(X)}\right)\psi(L^{-1}\mathbf{q}) .$$

From the condition

$$\mathcal{D}(L)\psi_r^{(0)}(\mathbf{q}) \stackrel{!}{=} w_r(k)e^{-\varepsilon_r ik_\mu x^\mu(\mathbf{q})} = w_r(k)e^{-\varepsilon_r ik_\mu v^\mu} , \tag{4.3}$$

we read of the spinors at $k \neq 0$ (using lemma D.5.13):

$$w_r(k) := \mathcal{D}_S\left(e^{\tau^{-1}(X)}\right)w_r(0) = \mathcal{D}_S\left(e^{\frac{1}{8}X_{\alpha\beta}[e^\alpha,e^\beta]}\right)w_r(0) =: S(L)\omega_r(0) .$$

Remark 4.6.5.
Comparing the exponentials yields:

$$\frac{mc}{\hbar}x^0((L^{-1})^\mu{}_\nu q^\nu \partial_\mu) = \frac{mc}{\hbar}(L^{-1})^0{}_\nu q^\nu = k_\nu q^\nu \quad \Rightarrow \quad (L^{-1})^0{}_\nu = \frac{\hbar}{mc}k_\nu .$$

To calculate $\mathcal{D}_S\left(e^{\frac{1}{8}X_{\alpha\beta}[e^\alpha,e^\beta]}\right)$ explicitly, one would have to express $e^{\frac{1}{8}X_{\alpha\beta}[e^\alpha,e^\beta]}$ in the Clifford basis. However, for the purpose of the textbook normalization, it is enough to observe, that the exponential map and the representation commute and thus:

$$\begin{aligned} S(L) = \mathcal{D}_S\left(e^{\frac{1}{8}X_{\alpha\beta}[e^\alpha,e^\beta]}\right) &= e^{\mathcal{D}_S(\frac{1}{8}X_{\alpha\beta}[e^\alpha,e^\beta])} \\ &= e^{\frac{1}{8}X_{\alpha\beta}[\mathcal{D}_S(e^\alpha),\mathcal{D}_S(e^\beta)]} = e^{\frac{1}{8}X_{\alpha\beta}[\gamma^\alpha,\gamma^\beta]} . \end{aligned} \tag{4.4}$$

Since $X_{\alpha\beta} \in \mathbb{R}$, it follows that

$$\begin{aligned} S(L)^\dagger &= e^{\frac{1}{8}X_{\alpha\beta}[\gamma^\alpha,\gamma^\beta]^\dagger} = e^{\frac{1}{8}X_{\alpha\beta}[(\gamma^\beta)^\gamma,(\gamma^\alpha)^\gamma]} = e^{\frac{1}{8}X_{\alpha\beta}\gamma^0[\gamma^\beta,\gamma^\alpha]\gamma^0} \\ &= e^{\gamma^0(-\frac{1}{8}X_{\alpha\beta}[\gamma^\alpha,\gamma^\beta])\gamma^0} = \gamma^0 e^{-\frac{1}{8}X_{\alpha\beta}[\gamma^\alpha,\gamma^\beta]}\gamma^0 = \gamma^0 S(L)^{-1}\gamma^0 \ . \end{aligned}$$

To obtain the second line, we have used, that

$$\left(\gamma^0[\gamma^\mu,\gamma^\nu]\gamma^0\right)^n = \gamma^0[\gamma^\mu,\gamma^\nu]\gamma^0\gamma^0[\gamma^\mu,\gamma^\nu]\dots[\gamma^\mu,\gamma^\nu]\gamma^0 = \gamma^0[\gamma^\mu,\gamma^\nu]^n\gamma^0 \ .$$

This result, together with the construction of $w_r(k)$ from the already normalized $w_r(0)$ leads to the following normalization:

Lemma 4.6.6.
It holds that

$$u_s(k)^\dagger\gamma^0 u_{s'}(k) = \delta_{ss'} \qquad \text{and} \qquad v_s(k)^\dagger\gamma^0 v_{s'}(k) = -\delta_{ss'} \ .$$

Proof 4.6.7.

$$\begin{aligned} w_r(k)^\dagger\gamma^0 w_{r'}(k) &= w_r(0)^\dagger S(L)^\dagger\gamma^0 S(L) w_{r'}(0) = w_r(0)^\dagger\gamma^0 S(L)^{-1}\gamma^0\gamma^0 S(L) w_{r'}(0) \\ &= w_r(0)^\dagger\gamma^0 w_{r'}(0) = \varepsilon_r\delta_{rr'} \ . \end{aligned}$$

□

This normalization condition is the reason for a notation in the literature, that is easily confused with complex conjugation:

Definition 4.6.8.
Let ψ be a spinor, then one defines

$$\bar{\psi} = \psi^\dagger\gamma^0 \ .$$

4.6.3. Lorentz boost of spinors

So far, the abstract rules for the action of $\mathrm{SO}(\mathbb{R}^4, g)$ on spinors have been enough to prove important properties. To find further identities for $u_s(k)$ and $v_s(k)$, an explicit expression for $S(L)$ is needed. To find such an explicit expression, the generator X of L has to be found first.

Lemma 4.6.9.
Let L be a Lorentz boost that boosts $(\frac{mc}{\hbar}, 0)$ to $(\frac{\omega(k)}{c}, k)$, where $\omega(k) = \sqrt{\left(\frac{mc^2}{\hbar}\right)^2 + \|k\|^2c^2}$. Then for

$$X^\mu{}_\nu = \frac{\theta(k)}{\|k\|}\sum_i k^i(\delta^0_\nu\delta^\mu_i + \delta^i_\nu\delta^\mu_0)$$

with $\theta(k) = \cosh^{-1}(\varphi(k))$ *and* $\varphi(k) = \sqrt{1 + \left(\frac{\|k\|\hbar}{mc}\right)^2}$

it holds that

Proof 4.6.10.
First we consider

$$Y^{\mu}{}_{\nu} = \frac{1}{\|k\|} \sum_i k^i (\delta^0_\nu \delta^\mu_i + \delta^i_\nu \delta^\mu_0) \ .$$

$$\Rightarrow \qquad Y = \begin{pmatrix} 0 & {}^{k^1}\!/_{\|k\|} & {}^{k^2}\!/_{\|k\|} & {}^{k^3}\!/_{\|k\|} \\ {}^{k^1}\!/_{\|k\|} & 0 & 0 & 0 \\ {}^{k^2}\!/_{\|k\|} & 0 & 0 & 0 \\ {}^{k^3}\!/_{\|k\|} & 0 & 0 & 0 \end{pmatrix}$$

We calculate:

$$\begin{aligned} Y^{\mu}{}_{\nu}\, Y^{\nu}{}_{\lambda} &= \frac{1}{\|k\|^2} \sum_{i,j=1}^{3} k^i k^j (\delta^0_\nu \delta^\mu_i + \delta^i_\nu \delta^\mu_0)(\delta^0_\lambda \delta^\nu_j + \delta^j_\lambda \delta^\nu_0) = \frac{1}{\|k\|^2} \sum_{i,j} k^i k^j (\delta^i_j \delta^\mu_0 \delta^0_\lambda \delta^j_\lambda \delta^\mu_i) \\ &= \frac{1}{\|k\|^2} \sum_{i,j} \left((k^i)^2 \delta^\mu_0 \delta^0_\lambda + k^i k^j \delta^\mu_i \delta^j_\lambda \right) = \delta^\mu_0 \delta^0_\lambda + \frac{1}{\|k\|^2} \sum_{i,j} k^i k^j \delta^\mu_i \delta^j_\lambda \ . \end{aligned}$$

$$\Rightarrow \qquad Y^2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & {}^{k^1k^1}\!/_{\|k\|^2} & {}^{k^1k^2}\!/_{\|k\|^2} & {}^{k^1k^3}\!/_{\|k\|^2} \\ 0 & {}^{k^2k^1}\!/_{\|k\|^2} & {}^{k^2k^2}\!/_{\|k\|^2} & {}^{k^2k^3}\!/_{\|k\|^2} \\ 0 & {}^{k^3k^1}\!/_{\|k\|^2} & {}^{k^3k^2}\!/_{\|k\|^2} & {}^{k^3k^3}\!/_{\|k\|^2} \end{pmatrix}$$

And also:

$$\begin{aligned} Y^{\mu}{}_{\nu}\, Y^{\nu}{}_{\lambda}\, Y^{\lambda}{}_{\alpha} &= \frac{1}{\|k\|} \sum_\ell k^\ell \left(\delta^0_\alpha \delta^\lambda_\ell + \delta^\ell_\alpha \delta^\lambda_0 \right) \left(\delta^\mu_0 \delta^0_\lambda + \frac{1}{\|k\|^2} \sum_{i,j} k^i k^j \delta^\mu_i \delta^j_\lambda \right) \\ &= \frac{1}{\|k\|} \sum_\ell k^\ell \left(\delta^\mu_0 \delta^\ell_\alpha + \frac{1}{\|k\|^2} \sum_i k^i k^\ell \delta^0_\alpha \delta^\mu_i \right) \\ &= \frac{1}{\|k\|} \sum_\ell k^\ell \delta^\ell_\alpha \delta^\mu_0 + \frac{1}{\|k\|} \sum_i k^i \delta^0_\alpha \delta^\mu_i \\ &= \frac{1}{\|k\|} \sum_\ell k^\ell (\delta^0_\alpha \delta^\mu_\ell + \delta^\ell_\alpha \delta^\mu_0) = Y^{\mu}{}_{\alpha} \ . \end{aligned}$$

Thus we see that $Y^3 = Y$ or in general:

$$Y^{2n} = Y^2 \qquad \text{and} \qquad Y^{2n+1} = Y \ .$$

It follows that

$$\exp(X) = \exp(\theta Y) = \sum_{n \in \mathbb{N}_0} \frac{\theta^n Y^n}{n!} = \mathbb{1} + \sum_{n \in 2\mathbb{N}} \frac{\theta^n Y^n}{n!} + \sum_{n \in 2\mathbb{N}_0 + 1} \frac{\theta^n Y^n}{n!}$$

$$
\begin{aligned}
&= \mathbb{1} + \sum_{n\in\mathbb{N}} \frac{\theta^{2n} Y^{2n}}{(2n)!} + \sum_{n\in\mathbb{N}_0} \frac{\theta^{2n+1} Y^{2n+1}}{(2n+1)!} \\
&= \mathbb{1} + \sum_{n\in\mathbb{N}} \frac{\theta^{2n}}{(2n)!} Y^2 + \sum_{n\in\mathbb{N}_0} \frac{\theta^{2n+1}}{(2n+1)!} Y \\
&= (\mathbb{1} - Y^2) + \cosh(\theta) Y^2 + \sinh(\theta) Y
\end{aligned}
$$

Now we have to show that $\exp(X)(\frac{mc}{\hbar}, 0) = (\frac{\omega(k)}{c}, k)$.

$$
\exp(X) \begin{pmatrix} \frac{mc}{\hbar} \\ 0 \end{pmatrix} = \begin{pmatrix} \cosh(\theta)\frac{mc}{\hbar} \\ \frac{k^i}{\|k\|} \sinh(\theta)\frac{mc}{\hbar} \end{pmatrix} .
$$

For the zeroth component we calculate:

$$
\begin{aligned}
\cosh(\theta)\frac{mc}{\hbar} &= \cosh(\cosh^{-1}(\varphi))\frac{mc}{\hbar} = \varphi\frac{mc}{\hbar} \\
&= \frac{mc}{\hbar}\sqrt{1 + \left(\frac{\|k\|\hbar}{mc}\right)^2} = \sqrt{\left(\frac{mc}{\hbar}\right)^2 + \|k\|^2} = \frac{\omega(k)}{c}
\end{aligned}
$$

For the i-th component one uses $\sinh(\cosh^{-1}(x)) = \sqrt{x^2 - 1}$:, to see that

$$
\frac{k^i}{\|k\|} \sinh(\theta)\frac{mc}{\hbar} = k^i .
$$

□

Remark 4.6.11.
For further calculations, it is useful to observe that

$$
\varphi(k) = \frac{E}{mc^2} = \frac{\sqrt{m^2c^4 + \hbar^2\|k\|^2c^2}}{mc^2}
$$

Lemma 4.6.12.
Let L, X and θ as in lemma 4.6.9, then it holds that

$$
S(L) \coloneqq \mathcal{D}_S(e^{\tau^{-1}(X)}) = \exp\left(\frac{\theta}{2}\frac{\sum_{i=1}^3 k^i\gamma^0\gamma^i}{\|k\|}\right) .
$$

Proof 4.6.13.
Becaus of (4.4), it only has to be shown that

$$
\frac{1}{8} X_{\alpha\beta}[\gamma^\alpha, \gamma^\beta] = \frac{\theta}{2}\frac{\sum_i k^i\gamma^0\gamma^i}{\|k\|}
$$

First we calculate:

$$
(X_{\alpha\beta}) = \theta \begin{pmatrix} 0 & {}^{k^1}/_{\|k\|} & {}^{k^2}/_{\|k\|} & {}^{k^3}/_{\|k\|} \\ -{}^{k^1}/_{\|k\|} & 0 & 0 & 0 \\ -{}^{k^2}/_{\|k\|} & 0 & 0 & 0 \\ -{}^{k^3}/_{\|k\|} & 0 & 0 & 0 \end{pmatrix}
$$

Then with $[\gamma^0, \gamma^i] = 2\gamma^0\gamma^i$

$$\begin{aligned}\frac{1}{8}X_{\alpha\beta}[\gamma^\alpha, \gamma^\beta] &= \frac{\theta}{8}\left(-\sum_{i=1}^{3}\frac{k^i}{\|k\|}[\gamma^i, \gamma^0] + \sum_{j=1}^{3}\frac{k^j}{\|k\|}[\gamma^0, \gamma^j]\right) \\ &= \frac{2\theta}{8}\sum_{i=1}^{3}\frac{k^i}{\|k\|}[\gamma^0, \gamma^i] = \frac{\theta}{2}\frac{\sum_{i=1}^{3}k^i\gamma^0\gamma^i}{\|k\|} \ .\end{aligned}$$

□

Theorem 4.6.14.
It holds that

$$S(L) = \sqrt{\frac{E+mc^2}{2mc^2}}\left(\mathbb{1} + \frac{\hbar c}{E+mc^2}\left(\sum_i k^i\gamma^0\gamma^i\right)\right) \ .$$

Proof 4.6.15.
First we calculate $(\sum_i k^i\gamma^0\gamma^i)^n$. Using $\gamma^i\gamma^j = -\gamma^j\gamma^i$ for $i \neq j$:

$$\begin{aligned}\left(\sum_i k^i\gamma^0\gamma^i\right)^2 &= +\sum_i (k^i)^2\mathbb{1} - k^1k^2(\gamma^1\gamma^2 + \gamma^2\gamma^1) - k^1k^3(\gamma^1\gamma^3 + \gamma^3\gamma^1) \\ &\quad - k^1k^2(\gamma^3\gamma^2 + \gamma^2\gamma^3) \\ &= \|k\|^2\mathbb{1} \ .\end{aligned}$$

$$\Rightarrow \qquad \left(\sum_i k^i\gamma^0\gamma^i\right)^{2n} = \|k\|^{2n}\mathbb{1}$$

$$\text{and} \quad \left(\sum_i k^i\gamma^0\gamma^i\right)^{2n+1} = \|k\|^{2n}\left(\sum_i k^i\gamma^0\gamma^i\right) \ .$$

This allows to calculate:

$$\begin{aligned}
\exp\left(\frac{\theta}{2}\frac{\sum_{i=1}^{3} k^i\gamma^0\gamma^i}{\|k\|}\right) &= \sum_{n\in\mathbb{N}_0} \frac{\left(\frac{\theta}{2}\frac{\sum_{i=1}^{3} k^i\gamma^0\gamma^i}{\|k\|}\right)^n}{n!} = \dots \\
&= \sum_{n\in\mathbb{N}_0} \frac{\left(\frac{\theta}{2\|k\|}\right)^{2n}}{(2n)!}\left(\sum_i k^i\gamma^0\gamma^i\right)^{2n} \\
&\qquad + \sum_{n\in\mathbb{N}_0} \frac{\left(\frac{\theta}{2\|k\|}\right)^{2n+1}}{(2n+1)!}\left(\sum_i k^i\gamma^0\gamma^i\right)^{2n+1} \\
&= \sum_{n\in\mathbb{N}_0} \frac{\left(\frac{\theta}{2}\right)^{2n}}{(2n)!}\mathbb{1} + \frac{1}{\|k\|}\sum_{n\in\mathbb{N}_0} \frac{\left(\frac{\theta}{2}\right)^{2n+1}}{(2n+1)!}\left(\sum_i k^i\gamma^0\gamma^i\right) \\
&= \cosh\left(\frac{\theta}{2}\right)\mathbb{1} + \frac{\sinh\left(\frac{\theta}{2}\right)}{\|k\|}\left(\sum_i k^i\gamma^0\gamma^i\right) \\
&= \cosh\left(\frac{\theta}{2}\right)\left(\mathbb{1} + \frac{\sinh\left(\frac{\theta}{2}\right)}{\cosh\left(\frac{\theta}{2}\right)\|k\|}\left(\sum_i k^i\gamma^0\gamma^i\right)\right)
\end{aligned}$$

Using the trigonometric identities $\cosh(\frac{1}{2}\cosh^{-1}(\varphi)) = \frac{\sqrt{\varphi+1}}{\sqrt{2}}$ we find:

$$\cosh\left(\frac{\theta}{2}\right) = \sqrt{\frac{E+mc^2}{2mc^2}}$$

Another trigonometric identity is $\tanh(\frac{\varphi}{2}) = \frac{\tanh(\varphi)}{1+\sqrt{1-\tanh(\varphi)^2}}$. Since we already have seen that $\cosh(\theta) = \varphi = \frac{E}{mc^2}$ and $\sinh(\theta) = \frac{\|k\|\hbar}{mc}$, we get $\tanh(\theta) = \frac{\|k\|\hbar}{E}$. Thus:

$$\tanh\left(\frac{\theta}{2}\right) = \frac{\frac{\|k\|\hbar c}{E}}{1+\sqrt{1-\frac{\|k\|^2\hbar^2c^2}{E^2}}} = \frac{\|k\|\hbar c}{E+\sqrt{E^2-\hbar^2\|k\|^2c^2}} = \frac{\|k\|\hbar c}{E-mc^2}\,.$$

Plugging in, we obtain:

$$S(L) = \sqrt{\frac{E+mc^2}{2mc^2}}\left(\mathbb{1} + \frac{\hbar c}{E+mc^2}\left(\sum_i k^i\gamma^0\gamma^i\right)\right)\,.$$

□

4.6.4. Further identities for $u_s(k)$ and $v_s(k)$ in Dirac representation

In theorem 4.6.14 we have found an explicit expression for $S(L)$. Using that $\gamma^0\gamma^i = \alpha_i$ and the Dirac representation (see remark 3.2.9), one obtains the matrix from [BD94, (3.7)]:

$$S(L) = \sqrt{\frac{E+mc^2}{2mc^2}}\begin{pmatrix} 1 & 0 & \frac{p^3c}{E+mc^2} & \frac{p^-c}{E+mc^2} \\ 0 & 1 & \frac{p^+c}{E+mc^2} & \frac{-p^3c}{E+mc^2} \\ \frac{p^3c}{E+mc^2} & \frac{p^-c}{E+mc^2} & 1 & 0 \\ \frac{p^+c}{E+mc^2} & \frac{-p^3c}{E+mc^2} & 0 & 1 \end{pmatrix}\,,$$

where $p^3 = \hbar k^3$ and $p^{\pm} = \hbar(k^1 \pm ik^2)$. With this explicit matrix representation, the following identities boil down to direct calculations:

$$\boxed{u_s(k)^\dagger u_{s'}(k) = \frac{E}{mc^2}\delta_{ss'} = v_s k^\dagger v_{s'}(k)}$$

$$\boxed{\sum_s u_s(k) \otimes \bar{u}_s(k) = \frac{1}{2}\left(\mathbb{1} + \frac{\hbar}{mc}\gamma^\mu k_\mu\right)}$$

$$\Rightarrow \qquad \sum_s u_s(k)^\alpha \bar{u}_s(k)_\beta = \frac{1}{2}\left(\mathbb{1} + \frac{\hbar}{mc}\not{k}\right)^\alpha{}_\beta$$

$$\boxed{\sum_s v_s(k) \otimes \bar{v}_s(k) = -\frac{1}{2}\left(\mathbb{1} - \frac{\hbar}{mc}\gamma^\mu k_\mu\right)}$$

$$\Rightarrow \qquad \sum_s v_s(k)^\alpha \bar{v}_s(k)_\beta = -\frac{1}{2}\left(\mathbb{1} - \frac{\hbar}{mc}\not{k}\right)^\alpha{}_\beta$$

$$\Rightarrow \qquad \boxed{\sum_s u_s(k) \otimes \bar{u}_s(k) - v_s(k) \otimes \bar{v}_s(k) = \mathbb{1}}$$

4.6.5. Dirac field with relativistic conventions

With the spinors constructed in this section, the Dirac field gets a new normalization coefficient:

$$\boxed{\Psi(\mathbf{x}) = \int \frac{dk^3}{(2\pi)^3}\sqrt{\frac{mc^2}{\hbar\omega(k)}}\sum_{s=\pm}\left(e^{-i\mathbf{kx}}u_s(k)c_{+,s}(k) + e^{i\mathbf{kx}}v_s(-k)c^\dagger_{-,-s}(k)\right)}$$

The new normalization coefficients appear because of the following: In the proof of the anti commutation relations of the Dirac field in lemma 4.4.6 , it was used that $u_s(k)^\dagger u_{s'}(k) = \delta_{ss'} = v_s k^\dagger v_{s'}(k)$. With the relativistic convention the line reads $u_s k^\dagger u_{s'}(k) = \frac{E}{mc^2}\delta_{ss'} = v_s(k)^\dagger v_{s'}(k)$, such that each u and v need an additional coefficient $\sqrt{\frac{mc^2}{E(k)}} = \sqrt{\frac{mc^2}{\hbar\omega(k)}}$.

Part II.

Quantum field theory

The search for a relativistic covariant quantum theory has lead to a reinterpretation of a single particle theory to a many particle theory. In the course of this reinterpretation, a new object, called field operator appeared. This is the starting point of quantum field theory, where one studies quantum fields and their emergence from classical field theories. Another quantization approach that is introduced in this part is the concept of functional integrals.

5
Canonical quantization

Morally, quantum mechanics as detailed theory of the microscopic scale has to give rise to classical mechanics for a large amount of particles in the meter scale. However, classical theories are well understood because of their ususal readiness for experiments. In contrast, it took years and still does to test even simple quantum mechanical hypotheses. Thus, most of the time, one tries to generalize classical theories to a quantum theory. This process is called quantization of a (field) theory. However, there is no unique, perfect quantization scheme. Instead, numerous quantizations have proven to be more or less efficient for different problems. One such quantization scheme is the **canonical quantization**, inspired by the historically first transition from classical mechanics to single particle quantum mechanics.

5.1. Reminder: Analytical mechanics

As preparation for the canonical field quantization, the basics of analytical mechanics are briefly summarized. This section is more or less inspired by [Zir15].

5.1.1. Lagrange functions

Let $(A, V, +)$ be an affine space. A classical Lagrange function L, for a system with f degrees of freedom is a map:

$$L\colon \underbrace{A \times \ldots \times A}_{f\text{-times}} \times \underbrace{A \times \ldots \times A}_{f\text{-times}} \times \mathbb{R} \longrightarrow \mathbb{R} \ ,$$

$$(q_1, \ldots, q_f, v_1, \ldots, v_f, t) \equiv (q, v, t) \longmapsto L(q, v, t) \ .$$

In the context of physics, $v \equiv v(t)$ describes the velocities and $q \equiv q(t)$ describes the positions, i.e. $v = \dot{q}$. Also, the Lagrange function is the difference between the kinetic energy T and the potential energy U: $L = T - U$.

The action of the Lagrange function is defined by $S(\Gamma) := \int_{t_1}^{t_2} (L \circ \gamma)(t)\ dt$, where $t \mapsto \gamma(t)$ is a parametrization of the curve Γ. The equations of motions follow from Hamilton's principle of stationary action, i.e. $\delta S(\Gamma, \Psi) = 0$ for all curves Ψ with $\psi(t_1) = 0 = \psi(t_2)$. With the methods of chapter G, the equations of motion, better known as *Euler-Lagrange equations* follow:

$$\boxed{\frac{\partial L(q, \dot{q}, t)}{\partial q_i} - \frac{d}{dt}\frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}_i} = 0 \ .}$$

5.1.2. Legendre transformation and Hamilton function

Let $f\colon X \to \mathbb{R}$ be a continuous function for a vector space X. This function is called convex, if

$$\forall x, y \in X\colon f(x + t(y - x)) \leq f(x) + t(f(y) - f(x)) \quad \forall t \in [0, 1] \ .$$

Graphically this means, that the value of the function of points on the straight line between x and y is less than the corresponding value of the line in $\mathbb{R}$ between $f(x)$ and $f(y)$.[1] In the case $X = \mathbb{R}$ it can be understood as the function being always below secant lines.

Definition 5.1.1.
The **Legendre transformation** $\mathfrak{L}$ maps convex functions $C^0(X)$ onto functions $C^0(X^*)$ by

$$(\mathfrak{L}f)(p) \coloneqq \sup_{x \in X} \{p(x) - f(x)\} \ .$$

Lemma 5.1.2.
Let $f\colon X \to \mathbb{R}$ be convex, then the function $F_p\colon X \to \mathbb{R}\ ,\ x \mapsto p(x) - f(x)$ is concave.

Proof 5.1.3.
Let $x, y \in X$ be arbitrary and $t \in [0, 1]$:

$$\begin{aligned} F_p(x + t(y - x)) &\overset{p \text{ linear}}{=} p(x) + t(p(y) - p(x)) - f(x + t(y - x)) \\ &\overset{f \text{ convex}}{\geq} p(x) + t(p(y) - p(x)) - f(x) - t(f(y) - f(x)) \\ &= p(x) - f(x) + t(p(y) - f(y) - p(x) + f(x)) \\ &= F_p(x) + t(F_p(y) - F_p(x)) \ . \end{aligned}$$

□

Theorem 5.1.4.
Let $f \in C^1(X)$, then the Legendre transformation is given by

$$(\mathfrak{L}f)(p) = p(h(p)) - f(h(p))$$

where $h\colon X^* \to X$ is defined as the inverse of the map $g\colon X \to X^*,\ x \mapsto (df)_x$.

Proof 5.1.5.
Define the function $F_p(x) = p(x) - f(x)$. By definition, the function $F_p(x)$ has a

[1] Usually, the condition is written differently as $f(ty + (1 - t)x) \leq tf(y) + (1 - t)f(x)$.

global maximum in x_p, such that

$$F_p(x_p) = \sup_{x\in X} F_p(x) = \sup_{x\in X}\{p(x) - f(x)\} = (\mathfrak{L}f)(p) \ .$$

Since X has no boundary points as vector space, the global maximum is also a local maximum, such that $(dF_p)_{x_p}(v) = 0 \ \forall v \in X$ determines x_p. Since F_p is concave (lemma 5.1.2), x_p is determined uniquely, such that we need no further condition like the Hessian etc.. We calculate[2]:

$$(dp)_x(v) = \frac{d}{ds}\bigg|_{s=0} p(x+sv) = \frac{d}{ds}\bigg|_{s=0} p(x) + s\cdot p(v) = p(v) \ .$$

Hence:

$$(dF_p)_{x_p}(v) = p(v) - (df)_{x_p}(v) = 0 \quad \Leftrightarrow \quad p = (df)_{x_p} = g(x_p)$$

$$\Leftrightarrow \quad x_p = h(p) \ .$$

Finally, plugging in the expression for x_p proves the claim:

$$(\mathfrak{L}f)(p) = F_p(x_p) = p(h(p)) - f(h(p)) \ .$$

□

Corollary 5.1.6.
The relation between p and x_p is given by

$$x_p = h(p) \qquad \text{and} \qquad p = (df)_{x_p} \ .$$

Let $\{e_i\}$ be an arbitrary basis of X and $\{\vartheta^i\}$ the dual basis of X^*. Then:

$$(\mathfrak{L}f)(p) = p_i h^i(p) - f(h(p)) \ ,$$

where we have used the summation convention and $p = p_i\vartheta^i$ and $h = h^i e_i$. Furthermore:

$$p = (df)_{x_p} = \frac{\partial f(x_p)}{\partial x^i}dx^i \qquad \Rightarrow \qquad \boxed{p_i = \frac{\partial f(x_p)}{\partial x^i}} \tag{5.1}$$

Theorem 5.1.7.
Let $f \in C^2(X)$ be a convex function, then the following equalities hold:

$$i)\ D_p\mathfrak{L}f = h(p) \qquad\qquad ii)\ (\mathfrak{L}(\mathfrak{L}f))(x) = f(x) \ .$$

The latter means, that $\mathfrak{L}$ is an involution, i.e. $\mathfrak{L}^2 = \mathbb{1}$.

[2]The isomorphism of X and X^{**} allows for the even shorter calculation $dp(v) = v(p) = p(v)$.

Proof 5.1.8.
First we calculate $D_p(p \circ h)$. Let $v \in X^*$, then:

$$\begin{aligned} D_p(p \circ h)(v) &= \frac{d}{ds}\Big|_{s=0} (p + sv)(h(p + sv)) \\ &= \frac{d}{ds}\Big|_{s=0} p(h(p + sv)) + sv(h(p + sv)) \\ &= p(D_p h(v)) + v(h(p)) \equiv (p \circ D_p h)(v) + (h(p))(v) \ . \end{aligned}$$

Thus we find the first claim:

$$\begin{aligned} D_p \mathfrak{L} f &= p \circ D_p h + h(p) - D_{h(p)} f \circ D_p h \\ &= p \circ D_p h + h(p) - g(h(p)) \circ D_p H \\ &= p \circ D_p h + h(p) - p \circ D_p H = h(p) \ . \end{aligned}$$

For the second claim we define $\widetilde{h} \colon X \to X^*$ as the inverse of $p \mapsto D_p \mathfrak{L} f$. With $D_p \mathfrak{L} f = h(p)$ it follows that $h \circ \widetilde{h} = \mathbb{1}$. Using the canonical isomorphism $X^{**} \cong X$, we know that $\widetilde{h}(x)(x) = x(\widetilde{h}(x))$

$$\begin{aligned} (\mathfrak{L}^2 f)(x) &= x(\widetilde{h}(x)) - (\mathfrak{L} f)(\widetilde{h}(x)) \\ &= x(\widetilde{h}(x)) - \widetilde{h}(x)(h(\widetilde{h}(x))) + f(h(\widetilde{h}(x))) \\ &= x(\widetilde{h}(x)) - \widetilde{h}(x)(x) + f(x) = f(x) \ . \end{aligned}$$

□

The Hamilton function is defined as Legendre transformation of a Lagrange function for fixed q and t, i.e. $f \colon \dot{q} \mapsto L(q, \dot{q}, t)$: With theorem 5.1.4 and corollary 5.1.6 this means:

$$H(q, p, t) = p_i \dot{q}^i - L(q, \dot{q}, t)\Big|_{\dot{q} = \dot{q}_p} \ ,$$

For the momentum (5.1) then reads

$$p_i = \frac{\partial L(q, \dot{q}_p, t)}{\partial \dot{q}^i} \ .$$

It is common practice to identify $\dot{q}$ and $\dot{q}_p$ notationally and simply write

$$p_i = \frac{\partial L(q, \dot{q}, t)}{\partial \dot{q}^i} \ .$$

The Euler-Lagrange equations become *Hamilton's equation*:

$$\boxed{\dot{p}_k = -\frac{\partial H}{\partial q} \qquad \text{and} \qquad \dot{q}_k = \frac{\partial H}{\partial p} \ .}$$

5.1.3. Extension to fields

The extension of Lagrangian mechanics for fields is covered in chapter G, here we will only adopt further physical conventions. The Lagrange function $L(\phi_J^I, \phi_{J,\mu}^I)$ is defined in terms of Lagrange densities (also called Lagrangian) $\mathcal{L}(\phi_J^I, \phi_{J,\mu}^I)$ by

$$L(\phi_J^I, \phi_{J,\mu}^I) = \int \mathcal{L}(\phi_J^I, \phi_{J,\mu}^I)\ dx^1 \dots dx^n\ .$$

The zeroth coordinate corresponds to the time and is given by $x^0 = ct$ in Minkowski-spacetime. Since the difference is only a factor c we will consider the Euclidean case and use x^0 and t interchangeably. The action is as usual defined by

$$S[\phi] = \int L(\phi_J^I, \phi_{J,\mu}^I)\ dx^0 = \int \mathcal{L}(\phi_J^I, \phi_{J,\mu}^I)\ dx^0 \dots dx^n\ ,$$

in accordance to chapter G. In said chapter, it is shown that:

$$\boxed{\frac{\delta S}{\delta \phi_J^I} = \frac{\partial \mathcal{L}(\phi_J^I, \phi_{J,\mu}^I)}{\partial \phi_J^I} - \partial_\mu \frac{\partial \mathcal{L}(\phi_J^I, \phi_{J,\mu}^I)}{\partial \phi_{J,\mu}^I}\ .}$$

Following [AS10, chapter 1], we define:

Definition 5.1.9.
Let $\mathcal{L}(\phi_j, \partial_i \phi_j)$ be a Lagrangian. The **canonical momenta** π_ℓ, conjugate to ϕ_ℓ is defined by

$$\pi_\ell := \frac{\partial \mathcal{L}(\phi_j, \partial_i \phi_j)}{\partial \dot{\phi}_\ell}\ .$$

If the relation between π_ℓ and $\dot{\phi}_\ell$ is invertible, i.e. $\dot{\phi}_\ell$ can be written in terms of the π_j: $\dot{\phi}_\ell(\phi_j, \pi_j)$, the **Hamilton density** $\mathcal{H}$ is defined by

$$\mathcal{H}(\phi_j, \pi_j) = \sum_\ell \pi_\ell \cdot \dot{\phi}_\ell - \mathcal{L}(\phi_j, \partial_i \phi_j)\ .$$

Lemma 5.1.10.
Let $L = \int \mathcal{L}\ dx^1 \dots dx^n$ be a Lagrange function, then the canonical momentum is given by

$$\pi_\ell(x) = \frac{\delta L}{\delta \dot{\phi}_\ell}\ .$$

Proof 5.1.11.
The Lagrange function defines a functional $L[\dot{\phi}_j]$ of $\dot{\phi}_j$ but not $\ddot{\phi}_j$. The Euler-Lagrange equations for $\dot{\phi}_j$ and $\ddot{\phi}_j$ (note that $\frac{\partial \mathcal{L}}{\partial \ddot{\phi}_\ell} = 0$) yield:

$$\delta L(\dot{\phi}_j, \psi_j) = \int \sum_\ell \frac{\partial \mathcal{L}}{\partial \dot{\phi}_\ell} \psi_j\ dx^1 \dots dx^n \overset{!}{=} \int \sum_\ell \frac{\delta L}{\delta \dot{\phi}_\ell} \psi_j\ dx^1 \dots dx^n$$

$$\Rightarrow \qquad \frac{\partial \mathcal{L}}{\partial \dot{\phi}_\ell} = \frac{\delta L}{\delta \dot{\phi}_\ell} \ .$$

□

Theorem 5.1.12.
Let $\mathcal{H}$ be a Hamilton density, derived as Legendre transform from a Lagrange density and $H = \int \mathcal{H}\, dx^1 \dots dx^n$ the Hamilton function. Then, the ***canonical equations of motion*** *read*

$$\dot{\phi}_\ell = \frac{\delta H}{\delta \pi_\ell} \qquad \text{and} \qquad \dot{\pi}_\ell = -\frac{\delta H}{\delta \phi_\ell} \ .$$

Proof 5.1.13.
The first equation is immediate, since H depends only on π_j but not $\dot{\pi}_j$, such that

$$\frac{\delta H}{\delta \pi_\ell} = \frac{\partial \mathcal{H}}{\partial \pi_\ell} = \dot{\phi}_\ell + \sum_k \pi_k \frac{\partial \dot{\phi}_k}{\partial \pi_\ell} - \sum_k \frac{\partial \mathcal{L}}{\partial \dot{\phi}_k} \frac{\partial \dot{\phi}_k}{\partial \pi_\ell} = \dot{\phi}_\ell \ ,$$

where $\frac{\partial \mathcal{L}}{\partial \dot{\phi}_k} = \pi_k$ was used. For $\frac{\delta H}{\delta \phi_\ell}$, the Euler-Lagrange equations yield:

$$\frac{\delta H}{\delta \phi_\ell} = \frac{\partial \mathcal{H}}{\partial \phi_\ell} - \partial_k \frac{\partial \mathcal{H}}{\partial (\partial_k \phi_\ell)} = -\frac{\partial \mathcal{L}}{\partial \phi_\ell} + \partial_t \pi_\ell + \partial_k \frac{\partial \mathcal{L}}{\partial (\partial_k \phi_\ell)} = \dot{\pi}_\ell - \frac{\delta L}{\delta \phi_\ell} \ .$$

Since we want to find solutions, that satisfy Hamilton's principle of least action, it holds that $\frac{\delta L}{\delta \phi_\ell} = 0$, which leads to the claim. □

5.1.4. Noether's theorem

Noether's theorem relating symmetries and continuity equations, that lead to conserved quantities with proper boundary conditions, also exists for fields. For this subsection we follow more or less [Wei14, section 1.3].

Remark 5.1.14.
This subsection should be read with caution, as the author is not sure, if the definitions do work in the general context. The goal is, to define symmetries and conservation laws, without the use of infinitesimals. So far, this does only work, if the Lagrangian does not explicitly depend on position/time.

Definition 5.1.15.
Let $\mathcal{L}(\phi_\nu, \partial_\mu \phi_\nu)$ be a Lagrangian density. A **continuous symmetry** is a differentiable one parameter group $(\phi_\nu)_s$ with $(\phi_\nu)_0 = \phi_\nu$, such that

$$\left.\frac{d}{ds}\right|_{s=0} \mathcal{L}((\phi_\nu)_s, \partial_\mu (\phi_\nu)_s) = \partial_\mu K^\mu \ ,$$

for some K^μ.

This definition of a symmetry allows to formulate Noether's theorem without the need of infinitesimals.

Theorem 5.1.16.
Let $(\phi_\nu)_s$ be a continuous symmetry with $\frac{d}{ds}\Big|_{s=0}(\phi_\nu)_s = X_\nu$ and $j^\mu = \sum_\nu \left(\frac{\partial \mathcal{L}}{\partial \partial_\mu \phi_\nu}\right) X_\nu - K^\mu$, then on-shell, i.e. if the ϕ satisfies the equations of motion, it holds that:

$$\partial_\mu j^\mu = 0 \ .$$

Proof 5.1.17.
Before using the assumption of any symmetry we can calculate (using $(\phi_\nu)_0 = \phi_\nu$):

$$\begin{aligned}
\frac{d}{ds}\Big|_{s=0} \mathcal{L}((\phi_\nu)_s, \partial_\mu(\phi_\nu)_s) &= \sum_\nu \frac{\partial \mathcal{L}}{\partial \phi_\nu} \frac{d}{ds}\Big|_{s=0} (\phi_\nu)_s + \sum_{\nu,\mu} \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi_\nu} \frac{d}{ds}\Big|_{s=0} \partial_\mu (\phi_\nu)_s \\
&= \sum_\nu \left(\frac{\partial \mathcal{L}}{\partial \phi_\nu} \frac{d}{ds}\Big|_{s=0} (\phi_\nu)_s + \sum_\mu \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \partial_\mu \phi_\nu} \frac{d}{ds}\Big|_{s=0} (\phi_\nu)_s \right) \right. \\
&\qquad \left. - \sum_\mu \left(\partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi_\nu} \right) \frac{d}{ds}\Big|_{s=0} (\phi_\nu)_s \right) \\
&= \sum_\nu \left(\sum_\mu \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \partial_\mu \phi_\nu} \frac{d}{ds}\Big|_{s=0} (\phi_\nu)_s \right) \right. \\
&\qquad \left. + \left(\frac{\partial \mathcal{L}}{\partial \phi_\nu} - \sum_\mu \left(\partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi_\nu} \right) \right) \frac{d}{ds}\Big|_{s=0} (\phi_\nu)_s \right) \\
&= \sum_\nu \left(\sum_\mu \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \partial_\mu \phi_\nu} \right) X_\nu + \frac{\delta L}{\delta \phi_\nu} X_\nu \right)
\end{aligned}$$

Since $(\phi_\nu)_s$ is a continuous symmetry it holds by definition, that:

$$\frac{d}{ds}\Big|_{s=0} \mathcal{L}((\phi_\nu)_s, \partial_\mu(\phi_\nu)_s) = \partial_\mu K^\mu \ .$$

Defining

$$j^\mu = \sum_\nu \left(\frac{\partial \mathcal{L}}{\partial \partial_\mu \phi_\nu} \right) X_\nu - K^\mu \ ,$$

and plugging in, yields:

$$\partial_\mu K^\mu = \partial_\mu j^\mu + \partial_\mu K^\mu + \frac{\delta L}{\delta \phi_\nu} X_\nu \qquad \Rightarrow \qquad \partial_\mu j^\mu = -\frac{\delta L}{\delta \phi_\nu} X_\nu \ .$$

If the field ϕ satisfies the equations of motion, then $\frac{\delta L}{\delta \phi_\nu} = 0$, which concludes the proof. □

Next we consider the case of translation invariance, which are known to correspond to energy (for x^0) and momentum (for x^j). Let $T(s,v)$ define the shift operator, defined by

$$(T(s,v)f)(x) = f(x+sv) \quad \text{with } s \in \mathbb{R}\ ,\ v \in \mathbb{R}^n\ .$$

Using a Taylor expansion, the action of the shift can be written as:

$$(T(s,v)f)(x) = f(p) + \partial_\rho f(x) \cdot s v^\rho + \mathcal{O}(s^2)\ .$$

Then, the family $(\phi_\nu)_s = T(s,v)\phi_\mu$ has the following parameter derivative:

$$\begin{aligned} X_\nu(x) &= \left.\frac{d}{ds}\right|_{s=0} (\phi_\nu)_s(x) = \left.\frac{d}{ds}\right|_{s=0} \phi_\nu(x+sv) \\ &= \left.\frac{d}{ds}\right|_{s=0} \Big(\phi_\nu(x) + s\cdot v^\rho \partial_\rho \phi_\nu(x) + \mathcal{O}(s^2)\Big) = v^\rho \partial_\rho \phi_\nu(x)\ . \end{aligned}$$

If the Lagrangian does not depend explicitly on x, i.e. $\mathcal{L}(x) = \mathcal{L}(\phi_\nu(x), \partial_\mu\phi_\nu(x))$, then

$$\left.\frac{d}{ds}\right|_{s=0} T(s,v)\mathcal{L} = v^\rho \partial_\rho \mathcal{L}\ .$$

Note, that the partial derivatives commute with shifts by constants (here sv) in the argument, such that $\partial_\mu T(s,v)f = T(s,v)\partial_\mu f$. Since the only position dependence arises from the fields, and partial derivatives commute, we observe that

$$\left.\frac{d}{ds}\right|_{s=0} \mathcal{L}(T(s,v)\phi_\nu(x), \partial_\mu T(s,v)\phi_\nu(x)) = \left.\frac{d}{ds}\right|_{s=0} T(s,v)\mathcal{L} = v^\rho\partial_\rho\mathcal{L} = \partial_\rho K^\rho\ .$$

Thus we see that $K^\mu = v^\mu\mathcal{L}$ and obtain:

$$j^\mu = \sum_\nu \left(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi_\nu}\right) v^\rho\partial_\rho\phi_\nu - v^\mu\mathcal{L} = v^\rho\left(\sum_\nu \left(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi_\nu}\right)\partial_\rho\phi_\nu - \delta^\mu_\rho\mathcal{L}\right) \equiv a^\rho T^\mu{}_\rho$$

$$\partial_\mu j^\mu = 0 \quad\Rightarrow\quad \partial_\mu T^\mu{}_\rho = 0 \quad\Rightarrow\quad \partial_\mu T^{\mu\nu} = 0\ .$$

The tensor with raised index $T^{\mu\nu}$ is called the **energy-momentum tensor**:

$$\boxed{T^{\mu\nu} = \sum_\eta \left(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi_\eta}\right)\partial^\nu\phi_\eta - g^{\mu\nu}\mathcal{L} \quad \text{and} \quad \partial_\mu T^{\mu\nu} = 0} \tag{5.2}$$

5.2. The quantization

The canonical quantization is a well covered topic in the literature. Here we briefly cover the idea, that can be found in [Sch08, section 12.3] for example.

To understand the motivation behind canonical quantization, it is helpful to recall the historical quantization. The initial theory is a classical mechanical theory, formulated in Hamiltonian mechanics. That is, there is a Hamilton function $H(p_j, q_k)$. Then the momenta p_j and positions q_k are promoted to hermitian operators $\widehat{p}_j$ and $\widehat{q}_k$, such that the following commutation relations are satisfied:

$$[\widehat{q}_k, \widehat{p}_j] = i\hbar\delta_{kj}\ , \qquad [\widehat{q}_j, \widehat{q}_k] = 0\ , \qquad [\widehat{p}_j, \widehat{p}_k] = 0\ .$$

In the case of a field theory, the Hamilton density $\mathcal{H}(\phi_j, \pi_j, x, t)$ defines the Hamilton function by $H = \int \mathcal{H}\, dV$. Here the "variables" of interest are the fields $\phi_j(x,t)$ and $\pi_j(x,t)$. The fields are promoted to operators ϕ and π, that have to satisfy the following equal time (anti) commutator relations for bosons (fermions):

$$[\phi_j(x,t), \phi_k(y,t)]_\pm = 0 \ , \qquad [\pi_j(x,t), \pi_k(y,t)]_\pm = 0 \ ,$$

$$[\phi_j(x,t), \pi_k(y,t)]_\pm = i\hbar\delta_{jk}\delta(x-y) \ ,$$

where $[\cdot,\cdot]_-$ is the usual commutator $[\cdot,\cdot]$ and $[\cdot,\cdot]_+$ is the usual anti commutator $\{\cdot,\cdot\}$.

Remark 5.2.1.
The Delta functions are used, since π_j and ϕ_j are momentum and position densities, such that

$$\int dx^3 \ [\phi_j(x,t), \pi_k(y,t)]_\pm = i\hbar\delta_{jk} \ .$$

To find expressions for the quantum fields, one can try to follow the harmonic oscillator analogy further, and try to bring the Hamilton density in the form $\mathcal{H}_p = \sum_j \frac{1}{2m}\pi_j^2 + \frac{m\omega^2}{2}\phi_j^2$, e.g. by Fourier transformation. Then, from the harmonic oscillator $H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2$, where $a = \sqrt{\frac{m\omega}{2\hbar}}(q + \frac{i}{m\omega}p)$, one defines:

$$a_j = \sqrt{\frac{m\omega}{2\hbar}}(\phi_j + \tfrac{i}{m\omega}\pi_j)$$

calculates $a_j^\dagger$ and expresses ϕ_j and π_j in terms of these creation/ annihilation operators. Finally, the meaning of these creation/ annihilation operators has to be found, e.g. as in subsection 5.4.2.

Another approach to obtain the quantum fields is to construct the field operators form second quantization, as we have done in section 4.4.

5.3. From particles to quantum fields

As a first application of the canonical quantization, we consider the $1d$ harmonic chain, as presented in [Zir14]. That is, a collection of N identical masses m connected by identical springs with spring constant c.

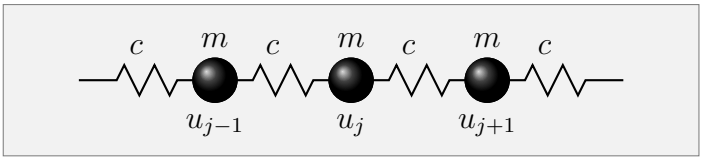


Figure 5.1.: Harmonic chain in one dimension.

This section also motivates the field theoretic description of systems, that are discrete or even finite.

5.3.1. Discrete equations of motion

Denoting the deviation from the equilibrium position of the j-th mass by u_j, a standard result from analytical mechanics is the Lagrange function:

$$L(u,\dot{u}) = \frac{m}{2}\sum_{j=1}^{N}\dot{u}_j - \frac{c}{2}\sum_{j=1}^{N}(u_j - u_{j-1})^2,$$

where we have used periodic boundary conditions ($u_0 = u_N$) for convenience. From the Euler-Lagrange equations, we find the equations of motion:

$$m\ddot{u}_j = -c(2u_j - u_{j+1} - u_{j-1}) \ .$$

To bring the equations of motion in a more suggestive form for later comparison with the continuum limit, we introduce the discrete $1d$ Laplace operator

$$\widetilde{\Delta}_{jj'} = -2\delta_{jj'} + \delta_{j(j'+1)} + \delta_{j(j'-1)} \qquad \Rightarrow \qquad (\widetilde{\Delta}u)_j = (-2u_j + u_{j+1} + u_{j-1})$$

Denoting $\sqrt{\frac{c}{m}} = \Omega$, the equations of motion become:

$$\frac{1}{\Omega^2}\ddot{u}_j - (\widetilde{\Delta}u)_j = 0 \ .$$

For fixed j', the eigen vectors of the discrete Laplace operator are

$$\psi^{j'} = e^{2\pi i \frac{jj'}{N}}$$

with eigen values

$$E_{j'} = -2 + e^{2\pi i\frac{j'}{N}} + e^{-2\pi i\frac{j'}{N}} = -2 + 2\cos(2\pi\tfrac{j'}{N}) = -4\sin^2(\pi\tfrac{j'}{N}) \ .$$

Thus the normal modes (principal solutions) for the equation of motion are:

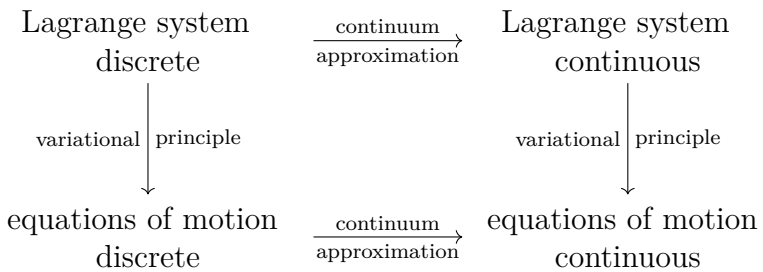
$$\psi^{j'}(t) = e^{i(2\pi\frac{jj'}{N} - \omega t)} \quad \text{with} \quad \omega = \Omega\sqrt{E_{j'}} = \Omega|2\sin(\pi\tfrac{j'}{N})|$$

The solutions can also be brought in a suggestive form, by labeling the components of the vectors with $x = j\frac{L}{N} \coloneqq ja$, where a is called *lattice constant*. Also, defining the wave number by $k = \frac{2\pi}{a}\frac{j'}{N}$, the normal modes can be written as

$$\psi(x,t) = e^{i(kx-\omega t)} \qquad \text{with} \qquad \omega = \Omega|2\sin(\tfrac{ka}{2})|$$

5.3.2. Continuum limit

Often it is easier to calculate in a continuous setting, using the results from analysis, instead of solving the original discrete setting. The idea/ hope is, that the following diagram commutes:



In the case of the harmonic chain, the continuum approximation amounts to the substitutions:

$$u_j(t) \longrightarrow \phi(x,t) \ , \qquad \dot{u}_j(t) \longrightarrow \dot{\phi}(x,t) \ , \qquad u_j(t) - u_{j-1}(t) \longrightarrow a\partial_x\phi(x,t) \ ,$$

$$\sum_{j=1}^{N} \longrightarrow \frac{1}{a}\int_0^L dx \ .$$

Hence, the Lagrange function of the harmonic chain becomes:

$$L(\phi,\dot{\phi},\phi') = \int_0^L \frac{m}{2a}\dot{\phi}^2 - \frac{ca}{2}\phi'^2 \ dx \ = \int_0^L \mathcal{L}(\phi,\dot{\phi},\phi') \ dx$$

Since $\mathcal{L}$ only depends on ϕ' and $\dot{\phi}$, but not ϕ, the variational principle reads:

$$\frac{\delta S}{\delta\phi} = \partial_t\frac{\partial\mathcal{L}}{\partial\dot{\phi}} + \partial_x\frac{\partial\mathcal{L}}{\partial\phi'} = 0 \ ,$$

where $S = \int_0^T L \ dt$. We find:

$$\partial_t\frac{\partial\mathcal{L}}{\partial\dot{\phi}} = \frac{m}{2a}\partial_t\partial_{\dot{\phi}}\dot{\phi}^2 = \frac{m}{a}\ddot{\phi}^2 \quad \text{and} \quad \partial_x\frac{\partial\mathcal{L}}{\partial\phi} = -\frac{ca}{2}\partial_x\partial_{\phi'}\phi'^2 = -ca\phi''$$

the equations of motion become the $1d$ wave equation with speed of sound $v_s = a\sqrt{\frac{c}{m}}$:

$$\frac{1}{v_s^2}\ddot{\phi} - \phi'' = 0 \ .$$

The normal modes of this partial differential equation are

$$\phi(x,t) = e^{i(kx-\omega t)} \qquad \text{with} \qquad \omega = v_s|k| \ .$$

Comparing the discrete with the continuous solution, we see that

	discrete	continuous
normal modes	$\psi(x,t) = e^{i(kx-\omega t)}$	$\phi(x,t) = e^{i(kx-\omega t)}$
ω	$\Omega\|2\sin(\frac{ka}{2})\| = \sqrt{\frac{c}{m}}\|2\sin(\frac{ka}{2})\|$	$v_s\|k\| = a\sqrt{\frac{c}{m}}\|k\|$

Noticing, that the first Taylor approximation of the discrete frequency ω is the continuous frequency, the continuum approximation works for small wave numbers k.

5.3.3. Scalar bosons and quantization

An application of the $1d$ harmonic chain is the usage as model for scalar bosons in $1d$. For the quantization the Hamilton function is needed, such that we calculate:

$$\pi = \frac{\delta L}{\delta\dot{\phi}} = \frac{\partial\mathcal{L}}{\partial\dot{\phi}} = \frac{m}{a}\dot{\phi} \quad \text{and} \quad \mathcal{L}(\phi,\pi) = \frac{a}{2m}\pi^2 - \frac{ca}{2}\phi'^2$$

$$\Rightarrow \qquad \mathcal{H} = \frac{a}{2m}\pi^2 + \frac{ca}{2}\phi'^2 \ .$$

To shorten the notation we write $\frac{a}{m} = \frac{1}{\mu}$ as mass density and $ca = \varkappa$ as elastic constant. Then, the Hamilton function reads:

$$H(\phi, \pi) = \int_0^L \frac{1}{2\mu}\pi^2 + \frac{\varkappa}{2}\phi'^2 \, dx$$

The next step is to use the Fourier transformation, or in this case, because of the compact interval $[0, L]$ the Fourier series. Because of the half interval $[0, L]$, one could develop the series only in real sine or cosine modes (Half range Fourier series). However, for computational ease, it is worthwhile to use the general complex series:

$$\phi(x) = \sqrt{\frac{1}{L}} \sum_{k \in \frac{2\pi}{L}\mathbb{Z}} \phi_k e^{ikx} \qquad \text{and} \qquad \pi(x) = \sqrt{\frac{1}{L}} \sum_{k \in \frac{2\pi}{L}\mathbb{Z}} \pi_k e^{ikx} \ ,$$

where we have used $k = \frac{2\pi n}{L}$ for the wave number. The Fourier coefficients are

$$\phi_k = \sqrt{\frac{1}{L}} \int_0^L e^{-ikx}\phi(x) \, dx \qquad \text{and} \qquad \pi_k = \sqrt{\frac{1}{L}} \int_0^L e^{-ikx}\pi(x) \, dx \ .$$

For the Fourier coefficients it holds that $(f')_k = ikf_k$ as in the Fourier transformation case (theorem E.1.6). We calculate, that there is a result, similar to lemma F.3.4, producing the Kronecker delta[3]:

$$\begin{aligned} \int_0^L \phi'^2 \, dx &= -\frac{1}{L} \int_0^L \sum_{k,\ell} k\ell\phi_k\phi_\ell e^{-i(k+\ell)x} \, dx = -\sum_{k,\ell} \delta_{k,-\ell} k\ell\phi_k\phi_\ell e^{-i(k+\ell)x} \\ &= \sum_k k^2\phi_k\phi_{-k} \ . \end{aligned}$$

In the same way one finds $\int_0^L \pi^2 \, dx = \sum_k \pi_k\pi_{-k}$. Also, noting that (after formally promoting the fields to operators)

$$\phi_k^\dagger := \overline{\phi_k} = \sqrt{\frac{1}{L}} \int_0^L \overline{e^{-ikx}\phi(x)} \, dx = \sqrt{\frac{1}{L}} \int_0^L e^{ikx}\phi(x) \, dx = \phi_{-k}$$

and similarly for π_k, so it follows that

$$\phi_k\phi_{-k} = |\phi_k|^2 \qquad \text{and} \qquad \pi_k\pi_{-k} = |\pi_k|^2$$

Plugging all in:

$$H = \sum_k \frac{|\pi_k|^2}{2\mu} + \frac{\varkappa k^2}{2}|\phi_k|^2 =: \sum_k H_k \ .$$

[3] In general:

$$\delta_{mn} = \frac{1}{2\pi} \int_0^{2\pi} e^{i(m-n)x} \, dx = \frac{1}{L} \int_0^L e^{\frac{2\pi i(m-n)x}{L} y} \, dy = \frac{1}{L} \int_0^L e^{i(k-\ell)y} \, dy \ .$$

Also note that $\delta_{mn} = \delta_{k\ell}$ for $k = \frac{2\pi n}{L}$ and $\ell = \frac{2\pi m}{L}$.

Following the general procedure, we define:

$$a_k = \sqrt{\frac{m\omega_k}{2\hbar}}(\phi_k + \tfrac{i}{m\omega_k}\pi_k) \qquad \Rightarrow \qquad a_k^\dagger\sqrt{\frac{m\omega_k}{2\hbar}}(\phi_{-k} - \tfrac{i}{m\omega_k}\pi_{-k}) \ ,$$

where we have set $\omega_k \coloneqq |k|\sqrt{\frac{\varkappa}{\mu}}$ and $m = \mu$. Solving for the fields, we obtain:

$$\phi_k = \sqrt{\frac{\hbar}{2m\omega_k}}\left(a_k + a_{-k}^\dagger\right) \qquad \text{and} \qquad \pi_k = -i\sqrt{\frac{\hbar m\omega_k}{2}}\left(a_k - a_{-k}^\dagger\right) \ .$$

For the field operators we find:

$$\boxed{\begin{aligned}\phi(x) &= \sqrt{\frac{1}{L}}\sum_{k\in\frac{2\pi}{L}\mathbb{Z}}\sqrt{\frac{\hbar}{2m\omega_k}}\left(e^{ikx}a_k + e^{-ikx}a_k^\dagger\right)\\ \pi(x) &= -i\sqrt{\frac{1}{L}}\sum_{k\in\frac{2\pi}{L}\mathbb{Z}}\sqrt{\frac{\hbar m\omega_k}{2}}\left(e^{ikx}a_k - e^{-ikx}a_k^\dagger\right) \ .\end{aligned}}$$

Using the exponential representation of the Kronecker delta, one can calculate the anti commutator rules for ϕ_k and π_k, from which the anti commutator relations of the creation/annihilation operators can be calculated:

$$\begin{aligned}[\phi_k, \pi_\ell] &= \frac{1}{L}\int_0^L\int_0^L e^{ikx}e^{i\ell y}[\phi(x), \pi(y)]\ dxdy\\ &= \frac{1}{L}\int_0^L\int_0^L e^{ikx}e^{i\ell y}i\hbar\delta(x-y)\ dxdy\\ &\frac{i\hbar}{L}\int_0^L e^{i(k-\ell)x}\ dx = i\hbar\delta_{k\ell} \ .\end{aligned}$$

The remaining anti commutators can be calculated the same way. We deduce, that the creation/annihilation operators satisfy:

$$[a_k, a_\ell^\dagger] = \delta_{k\ell} \qquad \text{and} \qquad [a_k, a_\ell] = 0 = [a_k^\dagger, a_\ell^\dagger] \ .$$

From $\phi_k^\dagger = \phi_{-k}$ it follows that $(a_k^\dagger)^\dagger = a_{-k}$, which is used to calculate:

$$H = \sum_{k\in\frac{2\pi}{L}\mathbb{Z}} H_k = \sum_{k\in\frac{2\pi}{L}\mathbb{Z}} \frac{\hbar m\omega_k}{2}\left(a_k a_k^\dagger + a_{-k}^\dagger a_{-k}\right)$$

$$\Rightarrow \qquad :H: = \sum_{k\in\frac{2\pi}{L}\mathbb{Z}} \frac{\hbar m\omega_k}{2}\left(a_k^\dagger a_k + a_{-k}^\dagger a_{-k}\right)$$

Using that ω_k is symmetric in k, one can calculate:

$$\boxed{:H: = \sum_{k\in\frac{2\pi}{L}\mathbb{Z}} \hbar m\omega_k\ a_k^\dagger a_k = \hbar\mu\sqrt{\frac{\varkappa}{\mu}}|k|\ a_k^\dagger a_k \ .}$$

5.4. The free real scalar field

In this section we will consider an example of the canonical quantization, where the fields are not confined in a one dimensional compact space region. To focus on the essential steps, natural units and real field will be used, unlike usual in quantum mechanics. This will become noticeable in the momentum operator for example.

This section follows [Wei14, Sections 1.4, 1.6 and 1.11].

5.4.1. Quantization

We consider the Klein-Gordon equation in natural units ($c = \hbar = 1$).

$$\left(\frac{\partial^2}{\partial_t^2} - \nabla^2 + m^2\right)\phi = 0 \ .$$

Since qft was introduced as relativistic description for quantum systems, we use the covariant formulation with metric $g = \mathrm{diag}(1, -1, -1, -1)$ and $x^0 = t$. Then the Klein-Gordon equation can be rewritten as

$$\left(\partial_\mu \partial^\mu + m^2\right)\phi = 0 \ .$$

The next step is to fin the Lagrangian (or directly the Hamiltonian). Unfortunately there is no unique way to find Lagrangians. The next best thing we can do, is to guess, and to check that the equations of motions are correct.

Lemma 5.4.1.
The Lagrangian $\mathcal{L}(\phi, \partial\phi, x) = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2$ yields the Klein-Gordon equation.

Proof 5.4.2.
All we need to do, is to calculate the Euler-Lagrange equations:

$$\frac{\partial}{\partial\phi}\mathcal{L} = -m^2\phi \ ,$$

$$\begin{aligned}\frac{\partial}{\partial(\partial_\mu\phi)}\mathcal{L} &= \partial_{x_\mu}|_{x_\mu=\partial_\mu\phi}\frac{1}{2}g^{\eta\nu}x_\eta x_\nu = \frac{1}{2}g^{\eta\nu}\delta_{\mu\eta}x_\nu + \frac{1}{2}g^{\eta\nu}\delta_{\mu\nu}x_\nu\Big|_{x_\mu=\partial_\mu\phi} \\ &= \frac{1}{2}(g^{\mu\nu}\partial_\nu\phi + g^{\eta^\mu}\partial_\eta\phi) = \partial^\mu\phi\end{aligned}$$

$$\Rightarrow \quad \partial_\mu\frac{\partial}{\partial(\partial_\mu\phi)}\mathcal{L} = \partial_\mu\partial^\mu\phi \ ,$$

$$\Rightarrow \qquad 0 = \frac{\delta L}{\delta\phi} = -\left(\partial_\mu\partial^\mu + m^2\right)\phi \quad \Leftrightarrow \quad \left(\partial_\mu\partial^\mu + m^2\right)\phi = 0 \ .$$

□

As the proof shows, one could also have chosen $-\mathcal{L}$, but would loose the convenient analogy to $L = T - V$. The next step is to find the canonical momentum π:

$$\pi = \frac{\delta L}{\delta \dot{\phi}} = \frac{\partial}{\partial \dot{\phi}} \mathcal{L} = \dot{\phi} \ .$$

Note that in the relativistic case $\dot{\phi} = \partial_0 \phi$ is meant here. In this case, the Hamiltonian density is rather simple to be found, because of $\dot{\phi}(\pi) = \pi$:

$$\mathcal{H}(\phi, \pi) = \pi^2 - \frac{1}{2}(\pi^2 + \partial_j \phi \partial^j \phi) + m^2 \phi^2 = \frac{\pi^2}{2} + \frac{1}{2}\left((\nabla\phi)^2 + m^2\phi^2\right) \ .$$

At this point, one formally applies the canonical quantization, here in the time-independent form. To develop further insight, we continue with a mode expansion, i.e. consider the momentum space ($p = k$ since $\hbar = 1$).

$$\phi(x) = \int \frac{dp^3}{(2\pi)^3} \ \widetilde{\phi}(p) e^{ipx} \qquad \text{and} \qquad \pi(x) = \int \frac{dp^3}{(2\pi)^3} \ \widetilde{\pi}(p) e^{ipx} \ .$$

In the momentum space, the Hamiltonian $H = \int dx^3 \ \mathcal{H}$ "decouples", i.e. becomes diagonal in the momentum. To see that, we calculate, using theorem E.1.6:

$$\begin{aligned}
(\nabla\phi)^2 &= \left(\nabla \int \frac{dp^3}{(2\pi)^3} \ \widetilde{\phi}(p) e^{ipx}\right)^2 \\
&= \sum_j \partial_j \left(\int \frac{dp^3}{(2\pi)^3} \ \widetilde{\phi}(p) e^{ipx}\right) \cdot \partial_j \left(\int \frac{dq^3}{(2\pi)^3} \ \widetilde{\phi}(q) e^{iqx}\right) \\
&= \sum_j \left(\int \frac{dp^3}{(2\pi)^3} \ ip_j \widetilde{\phi}(p) e^{ipx}\right) \left(\int \frac{dq^3}{(2\pi)^3} \ iq_j \widetilde{\phi}(q) e^{iqx}\right) \\
&= \int \frac{dp^3 dq^3}{(2\pi)^6} \ (-p \cdot q) e^{i(p+q)x} \widetilde{\phi}(p) \widetilde{\phi}(q) \ .
\end{aligned}$$

Considering $\int dx^3 \ (\nabla\phi)^2$ with lemma F.3.4 in the kernel notation:

$$\begin{aligned}
\int dx^3 \ (\nabla\phi)^2 &= \int dx^3 \ \int \frac{dp^3 dq^3}{(2\pi)^6} \ (-p \cdot q) e^{i(p+q)x} \widetilde{\phi}(p) \widetilde{\phi}(q) \\
&= \int \frac{dp^3 dq^3}{(2\pi)^6} \ (-p \cdot q) \widetilde{\phi}(p) \widetilde{\phi}(q) \ \int dx^3 \ e^{i(p+q)x} \\
&= \int \frac{dp^3 dq^3}{(2\pi)^6} \ (-p \cdot q) \widetilde{\phi}(p) \widetilde{\phi}(q) \cdot (2\pi)^3 \delta(p - (-q)) \\
&= \int \frac{dp^3}{(2\pi)^3} \ p^2 \widetilde{\phi}(p) \widetilde{\phi}(-p) \ .
\end{aligned}$$

In the same way, one can calculate that $\phi^2 = \int \frac{dp^3}{(2\pi)^3} \ \widetilde{\phi}(p) \widetilde{\phi}(-p)$. Since the scalar field in the position space is real, we obtain self adjoint operators, and it holds that

$$\phi^\dagger(p) \coloneqq \overline{\widetilde{\phi}(p)} = \int dx^3 \ \overline{e^{-ipx} \phi(x)} = \int dx^3 \ e^{-i(-p)x} \phi(x) = \widetilde{\phi}(-p) \ ,$$

such that $\widetilde{\phi}(p)\widetilde{\phi}(-p) = |\widetilde{\phi}(p)|^2$. Thus the Hamiltonian becomes:

$$\begin{aligned} H &= \int \frac{dp^3}{(2\pi)^3}\ \frac{1}{2}|\widetilde{\pi}(p)|^2 + \frac{1}{2}\left(p^2+m^2\right)|\widetilde{\phi}(p)|^2 \\ &\equiv \int \frac{dp^3}{(2\pi)^3}\ \frac{1}{2}|\widetilde{\pi}(p)|^2 + \frac{1}{2}\omega_p^2|\widetilde{\phi}(p)|^2 \equiv \int \frac{dp^3}{(2\pi)^3}\ \widetilde{\mathcal{H}}_p \end{aligned}$$

where we have defined $\omega_p = \sqrt{p^2+m^2}$ and $\widetilde{\mathcal{H}}_p = \frac{1}{2}|\widetilde{\pi}(p)|^2 + \frac{1}{2}\omega_p^2|\widetilde{\phi}(p)|^2$. Inspired by the harmonic oscillator, we now define:

$$a(p) = \sqrt{\frac{\omega_p}{2}}\left(\widetilde{\phi}(p) + \frac{i}{\omega}\widetilde{\pi}(p)\right) \quad \Rightarrow \quad a^\dagger(p) = \sqrt{\frac{\omega_p}{2}}\left(\widetilde{\phi}(-p) - \frac{i}{\omega}\widetilde{\pi}(-p)\right)\ .$$

Solving for $\widetilde{\phi}(p)$ and $\widetilde{\pi}(p)$, as is done for the harmonic oscillator, we find:

$$\widetilde{\phi}(p) = \sqrt{\frac{1}{2\omega_p}}\left(a(p) + a^\dagger(-p)\right) \qquad \text{and} \qquad \widetilde{\pi}(p) = -i\sqrt{\frac{\omega_p}{2}}\left(a(p) - a^\dagger(-p)\right)\ . \tag{5.3}$$

Plugging these into the Fourier-transformation for $\phi(x)$ and $\pi(x)$ and using

$$\int dp^n\ f(-p)e^{ipx} = \int dp^n\ f(p)e^{-ipx}\ ,$$

we obtain the mode expansion of the free field operators:

$$\boxed{\begin{aligned} \phi(x) &= \int \frac{dp^3}{(2\pi)^3}\ \sqrt{\frac{1}{2\omega_p}}\ a(p)e^{ipx} + a^\dagger(p)e^{-ipx} \\ \pi(x) &= -i\int \frac{dp^3}{(2\pi)^3}\ \sqrt{\frac{\omega_p}{2}}\ a(p)e^{ipx} - a^\dagger(p)e^{-ipx}\ . \end{aligned}}$$

From the commutation relations for $\phi(x)$ and $\pi(x)$ we calculate:

$$\begin{aligned} [\widetilde{\phi}(p), \widetilde{\pi}(q)] &= \int dx^3\ dy^3\ e^{-ipx}e^{-iqy}[\phi(x), \pi(y)] \\ &= \int dx^3\ dy^3\ e^{-ipx}e^{-iqy}i\delta(x-y) = i\int dx^3 e^{-i(p+q)x} \\ &= i(2\pi)^3\delta(p+q)\ . \end{aligned}$$

In the same way, we see that $[\widetilde{\phi}(p), \widetilde{\phi}(q)] = 0 = [\widetilde{\pi}(p), \widetilde{\pi}(q)]$. From that and (5.3), we deduce:

$$[a(p), a^\dagger(q)] = (2\pi)^3\delta(p-q) \qquad [a^\dagger(p), a^\dagger(q)] = 0 = [a(p), a(q)]\ .$$

These commutator relations show, that $a^\dagger(p)$ and $a(p)$ are of bosonic nature. If we had choosen $\{\cdot,\cdot\}$ for the field quantization, we would have got fermionic operators instead.

The last step is to calculate the normal ordered Hamiltonian in terms of the creation/annihilation operators. We notice from (5.3) and $\widetilde{\phi}^\dagger(p) = \widetilde{\phi}(-p)$, that $(a^\dagger(p))^\dagger = a(-p)$, such that:

$$\widetilde{\mathcal{H}}_p = \frac{1}{2}|\widetilde{\pi}(p)|^2 + \frac{1}{2}\omega_p^2|\widetilde{\phi}(p)|^2$$

$$
\begin{aligned}
&= \frac{1}{2}(-i)^2\frac{\omega_p}{2}\Big(a(p)-a^\dagger(-p)\Big)\Big(a(p)-a^\dagger(-p)\Big)^\dagger \\
&\qquad + \frac{\omega_p^2}{2}\frac{1}{2\omega_p}\Big(a(p)+a^\dagger(-p)\Big)\Big(a(p)+a^\dagger(-p)\Big)^\dagger \\
&= -\frac{\omega_p}{4}\Big(a(p)-a^\dagger(-p)\Big)\Big(a^\dagger(-p)-a(p)\Big) \\
&\qquad + \frac{\omega_p}{4}\Big(a(p)+a^\dagger(-p)\Big)\Big(a^\dagger(-p)+a(p)\Big) \\
&= \ldots = \frac{\omega_p}{2}\Big(a(p)a^\dagger(p)+a^\dagger(-p)a(-p)\Big) \\
\Rightarrow \qquad &:\mathcal{H}_p: = \frac{\omega_p}{2}\Big(a^\dagger(p)a(p)+a^\dagger(-p)a(-p)\Big)\ .
\end{aligned}
$$

Since $\omega_p = \sqrt{p^2+m^2}$ is symmetric in p we find:

$$\boxed{:H: = \int \frac{dp^3}{(2\pi)^3}\ \omega_p\ a^\dagger(p)a(p)\ .}$$

5.4.2. Momenta exitation

We want to find an interpretation for the $a(p)$ and $a^\dagger(p)$. It will turn out that these operators create momenta, such that it is pertinent to consider the momentum operator $:P:$. Recalling the energy momentum tensor, we are looking for T^{0j}, given by equation (5.2). We calculate:

$$\frac{\partial\mathcal{L}}{\partial(\partial_0\phi)} = \partial^\mu\phi \quad \Rightarrow \quad T^{0j} = \frac{\partial\mathcal{L}}{\partial(\partial_0\phi)}\partial^j\phi - g^{0j}\mathcal{L} = \dot\phi\partial^j\phi\ .$$

The momentum is then given by $p^j = \int dx^3\ \dot\phi\partial^j\phi$. Inserting the mode expanded operators, where $\dot\phi(\pi) = \pi$, we find after normal ordering:

$$P^j = \int \frac{dp^3}{(2\pi)^3}\ p^j a^\dagger(p)a(p)\ .$$

Summarizing the Hamilton operator and momentum operator to the four momentum $\mathfrak{P} = (H, P^1, \ldots, P^3)$, a direct calculation shows that:

$$[\mathfrak{P}^\mu, a^\dagger(p)] = p^\mu a^\dagger(p) \quad \text{and} \quad [\mathfrak{P}^\mu, a(p)] = -p^\mu a(p)\ ,$$

with $p^0 = \omega_p$. Since the operators $\mathfrak{P}^\mu$ are hermitian, there exist (pseudo) eigen states $|k^\mu\rangle$ with real eigen values k^μ, such that:

$$\mathfrak{P}^\mu|k^\mu\rangle = k^\mu|k^\mu\rangle\ .$$

From the commutator relations we see that:

$$\mathfrak{P}^\mu a^\dagger(p)|k^\mu\rangle = \Big(a^\dagger(p)\mathfrak{P}^\mu + p^\mu a^\dagger(p)\Big)|k^\mu\rangle = (k^\mu+p^\mu)|k^\mu\rangle\ .$$

In the same way:

$$\mathfrak{P}^\mu a(p)|k^\mu\rangle = (k^\mu - p^\mu)|k^\mu\rangle\ .$$

This means, that $a^\dagger(p)$ excites the state $|k^\mu\rangle$ with the momentum p^μ whereas $a(q)$ decreases the momentum by p^μ.

6 Functional integrals

In its best known form, integrals take region in space which is evaluated by an expression, i.e. a function or differential form. Functional integrals however take regions in the set of functions. This chapter does not focus on the theory of functional integrals in general, but introduces the important concepts for quantum field theory.

6.1. Feynman path integral

A prime example for functional integrals in physics is the Feynman path integral, whose interpretation leads to the saying, a quantum particle does not take one, but all paths. This section is mostly based on [Zir14, chapters I and II] and [AS10, chapters 1 and 3].

6.1.1. Construction of the path integral

We consider an autonomous system, described by the Hamilton operator H. Autonomous means, that H does not depend explicitly on the time. In section 1.5, we have already used the time evolution operator U_t. Here we employ the slightly more general form $U(t, t_0)$, defined by

$$i\hbar \frac{d}{dt} U(t, t_0) = H U(t, t_0) \qquad \text{and} \qquad |\Psi(t)\rangle = U(t, t_0)|\Psi(t_0)\rangle \ .$$

From the definition it is clear, that the time evolution operator maps a state $|\Psi(t_0)\rangle$ at time t_0 to the state $|\Psi(t)\rangle$ at time t. This means, the operator $U(t, t_0)$ takes states at time t_0 and has them evolve naturally under the Hamiltonian H until the time t, hence the name. For autonomous systems, the time evolution operator can be written explicitly as

$$U(t, t_0) = e^{-\frac{i}{\hbar} H \cdot (t - t_0)} \ .$$

We pass to the position representation and define:

Definition 6.1.1.
If there is a "function" $K(q', q, t', t)$, such that for all q, q' and t, t'

$$\Psi(q', t') = U(t', t)\Psi(q', t) = \int dq^n \ K(q', q, t', t)\Psi(q, t)$$

holds, it is called **propagator** of the system.

The word function in the definition is to be understood in a broader sense, as things like delta-functions are usually included. The motivation behind this definition, is the

convention of the physical literature (also discussed at the end of section F.3) to write also singular distributions G as integration-operator with integral kernel g:

$$G[\varphi] = \int dq^n \; g(q)\varphi(q) \; .$$

A better motivation fo such a definition might be the suggestiveness of Dirac notation together with the pseudo completeness relation $\int dq^n \; |q\rangle\langle q| = \mathbb{1}$:

$$\begin{aligned}\Psi(q',t') = \langle q'|\Psi(t')\rangle = \langle q'|U(t',t)|\Psi(t)\rangle &= \int dq^n \; \langle q'|U(t',t)|q\rangle\langle q|\Psi(t)\rangle \\ &\equiv \int dq^n \; K(q',q,t',t)\Psi(q,t) \; .\end{aligned}$$

In the physical literature, the propagator $\langle q'|U(t',t)|q\rangle$ is understood as the (q',q)-matrix elements. However, formally one should be careful with such terms, as a Hilbert basis needs to be countable, whereas $|q\rangle$ is not.

The spacial case $U_t \coloneqq U(t,0)$ reveals further structure of the set $\{U_t\}_{t\in\mathbb{R}}$. This set is a one-parameter group. Indeed, the explicit form $U_t = e^{-\frac{i}{\hbar}Ht}$ allows to easily see, that $U_t \circ U_s = U_{t+s}$. Another important result follows from the observation, that the action of $U(t,t_0)$ on $|\Psi(t_0)\rangle$ is independent of the argument of $|\Psi\rangle$, since $U(t,t_0)\colon \mathcal{H} \to \mathcal{H}$, where $\mathcal{H}$ denotes the Hilbert space. This allows to see, that:

$$U(t,t_0)|\Psi(t_0)\rangle = e^{-\frac{i}{\hbar}H\cdot(t-t_0)}|\Psi(t_0)\rangle = U_{t-t_0}|\Psi(t_0)\rangle$$

$$\Rightarrow \qquad U(t,t_0) = U_{t-t_0} \in \{U_t\}_{t\in\mathbb{R}} \; .$$

Hence, the problem can be reduced to $\{U_t\}_{t\in\mathbb{R}}$, which offers helpful structure as one-parameter group. In terms of propagators, the group structure becomes (using Fubini-Tonelli):

$$\begin{aligned}\Psi(q,t+s) = \langle q|U_{t+s}|\Psi\rangle &= \langle q|U_t \circ U_s|\Psi\rangle \\ &= \langle q|U_t|U_s\Psi\rangle = \int dx^n \langle q|U_t|x\rangle\langle x|U_s|\Psi\rangle \\ &= \int dx^n \; \langle q|U_t|x\rangle \int dy^n \; \langle x|U_s|y\rangle\langle y|\Psi\rangle \\ &\overset{F.T.}{=} \int dy^n \; \left(\int dx^n \; \langle q|U_t|x\rangle\langle x|U_s|y\rangle\right)\langle y|\Psi\rangle \; .\end{aligned}$$

From the formula above we read off, in accordance to the pseudo completeness relation:

$$\langle q|U_{t+s}|q'\rangle = \int dy^n \; \langle q|U_t|y\rangle\langle y|U_s|q'\rangle \; .$$

Also the calculation shows, how the kernel representation works for the composition of operators in general:

$$\langle q|A \circ B|q'\rangle = \int dy^n \; \langle q|A|y\rangle\langle y|B|q'\rangle \; .$$

Remark 6.1.2.
The concept of the Feynman path integral is to split the time evolution operator U_T in a sequence of propagators $U_{t_1} \circ \ldots \circ U_{t_N}$ an then to take the limit $N \to \infty$. In terms of propagators, this will result in a bunch of integrations, that will be called path integral.

For the Feynman path integral, at least in this introduction, the simple form $H = T+V$ is used for the Hamilton operator. For the time development operator $U_{\Delta t}$ in the limit of $\Delta t \to 0$ we want to find a product of operators, that are more handleable. First it holds that:

$$U_{\Delta t} = e^{-\frac{i}{\hbar}H\Delta t} = e^{-\frac{i}{\hbar}T\Delta t - \frac{i}{\hbar}V\Delta t} \ .$$

On the other hand, using theorem B.1.3 it follows that:

$$e^{\frac{\varepsilon}{2}B}e^{\varepsilon A}e^{\frac{\varepsilon}{2}B} = e^{\frac{\varepsilon}{2}B}\left(e^{\varepsilon A + \frac{\varepsilon}{2}B + \frac{\varepsilon^2}{4}[A,B] + \mathcal{O}(\varepsilon^3)}\right) = e^{\varepsilon(A+B)+\mathcal{O}(\varepsilon^3)} \ .$$

Hence for $\varepsilon = -\frac{i\Delta t}{\hbar}$:

$$e^{-\frac{i\Delta t}{2\hbar}V}e^{-\frac{i\Delta t}{\hbar}T}e^{-\frac{i\Delta t}{2\hbar}V} = e^{-\frac{i\Delta t}{\hbar}(T+V)+\mathcal{O}(\Delta t^3)} = e^{-\frac{i\Delta t}{\hbar}H+\mathcal{O}(\Delta t^3)} \ .$$

In the case of bounded operators T and V, the $\mathcal{O}(\Delta t^3)$ terms could be ignored for $\Delta t \to 0$, also for the first derivative, that generates the Schrödinger equation. Yet T is not bounded here and for most cases, neither is V. Still, the assumption is, that

$$e^{-\frac{i\Delta t}{\hbar}H} \simeq e^{-\frac{i\Delta t}{2\hbar}V}e^{-\frac{i\Delta t}{\hbar}T}e^{-\frac{i\Delta t}{2\hbar}V}$$

holds for $\Delta t \to 0$ (at least for reasonable physical states). To utilize this assumption, that separates the potential and kinetic time evolution, we calculate the propagator (with the help of the pseudo completeness relation $\int \frac{dp^n}{(2\pi\hbar)^n}\ |p\rangle\langle p|$ because $|p\rangle$ are eigenstates of the momentum operator $\mathbf{p}$)[1]:

$$\begin{aligned}
\langle x|e^{-\frac{i\Delta t}{2m\hbar}\mathbf{p}^2}|y\rangle &= \frac{1}{(2\pi\hbar)^n}\int dp^n\ \langle x|e^{-\frac{i\Delta t}{2m\hbar}\mathbf{p}^2}|p\rangle\langle p|y\rangle \\
&= \frac{1}{(2\pi\hbar)^n}\int dp^n\ e^{-\frac{i\Delta t}{2m\hbar}p^2}\langle x|p\rangle\langle p|y\rangle \\
&= \frac{1}{(2\pi\hbar)^n}\int dp^n\ e^{-\frac{i\Delta t}{2m\hbar}p^2}e^{\frac{i}{\hbar}x\cdot p}e^{-\frac{i}{\hbar}y\cdot p} \\
&= \frac{1}{(2\pi\hbar)^n}\int dp^n e^{-\frac{i\Delta t}{2m\hbar}p^2 - \frac{i}{\hbar}(y-x)\cdot p} \\
&= \frac{1}{(2\pi\hbar)^n}\left(\sqrt{\frac{\pi}{\frac{i\Delta t}{2m\hbar}}}\right)^n \exp\left(-\frac{(y-x)^2}{4\hbar\cdot\frac{i\Delta t}{2m\hbar}}\right) \\
&= \left(\frac{m}{2\pi i\hbar\Delta t}\right)^{\frac{n}{2}}\ e^{\frac{im}{2\Delta t}(y-x)^2} \ .
\end{aligned}$$

[1] Although there do exist methods to make these completeness relations rigorous, one has to be careful of the domain. However, as there are several instances of sloppiness in this section, we do not bother to proof the result any further.

To obtain the last line, corollary 6.1.5 was used. The potential kernel is much easier to find, since $|y\rangle$ is already an eigenstate of $\mathbf{y}$:

$$\langle x|e^{-\frac{i\Delta t}{2\hbar}V(\mathbf{y})}|y\rangle = e^{-\frac{i\Delta t}{2\hbar}V(y)}\langle x|y\rangle = e^{-\frac{i\Delta t}{2\hbar}V(y)}\delta(x-y)\ .$$

In the $\Delta t \to 0$ limit, the full propagator becomes:

$$\begin{aligned}
\langle x|e^{-\frac{i\Delta t}{\hbar}H}|y\rangle &\simeq \langle x|e^{-\frac{i\Delta t}{2\hbar}V}e^{-\frac{i\Delta t}{\hbar}T}e^{-\frac{i\Delta t}{2\hbar}V}|y\rangle \\
&= \int dq^n \int dq'^n\ \langle x|e^{-\frac{i\Delta t}{2\hbar}V}|q\rangle\langle q|e^{-\frac{i\Delta t}{\hbar}T}|q'\rangle\langle q'|e^{-\frac{i\Delta t}{2\hbar}V}|y\rangle \\
&= \left(\frac{m}{2\pi i\hbar\Delta t}\right)^{\frac{n}{2}} \int dq^n \int dq'^n\ e^{-\frac{i\Delta t}{2\hbar}V(q)}e^{\frac{im}{2\Delta t}(q'-q)^2}e^{-\frac{i\Delta t}{2\hbar}V(q')} \\
&\qquad\qquad \cdot\,\delta(q-x)\delta(q'-y) \\
&= \left(\frac{m}{2\pi i\hbar\Delta t}\right)^{\frac{n}{2}} e^{-\frac{i\Delta t}{2\hbar}V(x)}e^{\frac{im}{2\Delta t}(y-x)^2}e^{-\frac{i\Delta t}{2\hbar}V(y)} \\
&= \left(\frac{m}{2\pi i\hbar\Delta t}\right)^{\frac{n}{2}}\ \exp\left(\frac{i\Delta t}{\hbar}\left(\frac{m}{2}\frac{(y-x)^2}{\Delta t^2} - \frac{1}{2}(V(x)+V(y))\right)\right)
\end{aligned}$$

With these preparations done, the announced concept of the Feynman path integral can be realized. That is, we intend to compute the kernel $\langle x|e^{-\frac{it}{\hbar}H}|y\rangle$ by discretizing the interval $[0,t]$, setting $\Delta t = \frac{t}{N}$ and taking the limit $N \to \infty$:

$$\begin{aligned}
\langle x|e^{-\frac{it}{\hbar}H}|y\rangle = \lim_{N\to\infty}\left(\frac{m}{2\pi i\hbar\Delta t}\right)^{\frac{N\cdot n}{2}} \int_{\mathbb{R}^n} dq_1^n \ldots \int_{\mathbb{R}^n} dq_{N-1}^n \\
\exp\left(\frac{i\Delta t}{\hbar}\sum_{j=0}^{N-1}\left(\frac{m}{2}\frac{(q_{j+1}-q_j)^2}{\Delta t^2} - \frac{1}{2}(V(q_{j+1})+V(q_j))\right)\right)\Bigg|_{\substack{q_N=x\\ q_0=y}}\ .
\end{aligned}$$

The right hand side is called **Feynman path integral**. There are several assumptions that are involved and that follow. Most importantly, it is assumed without proof (here), that the limit exists, though there is a factor $N^{\frac{N}{2}}$ in front, due to $\Delta t = \frac{t}{N}$. The next assumption is, that

$$\begin{aligned}
&\frac{i\Delta t}{\hbar}\sum_{j=0}^{N-1}\left(\frac{m}{2}\frac{(q_{j+1}-q_j)^2}{\Delta t^2} - \frac{1}{2}(V(q_{j+1})+V(q_j))\right) \\
&\quad\xrightarrow{\Delta t\to 0} \frac{i}{\hbar}\int_0^t \frac{m}{2}\dot{q}(\tau)^2 - V(q(\tau))\ d\tau \equiv \frac{i}{\hbar}\int_0^t L(q(\tau))\ d\tau = \frac{i}{\hbar}S[q(t)]\ ,
\end{aligned}$$

of course with the boundary conditions $q(0) = y$ and $q(t) = x$. To obtain the notation of the literature, the product of infinitely man integrals (together with the coefficients in front) gets denoted by $\int D[q(t)]$:

$$\boxed{\langle x|e^{-\frac{it}{\hbar}H}|y\rangle = \int D[q(t)]\ e^{\frac{i}{\hbar}S[q(t)]}\ .}$$

6.1.2. Interpretation of the path integral

To understand the meaning of the path integral, in the sense to give a physical intuition, it is essential to understand the composition of propagators. For that reason we consider again:

$$\langle x|U_{s+\tau}|y\rangle = \int_{\mathbb{R}} dq\ \langle x|U_s|q\rangle\langle q|U_\tau|y\rangle\ .$$

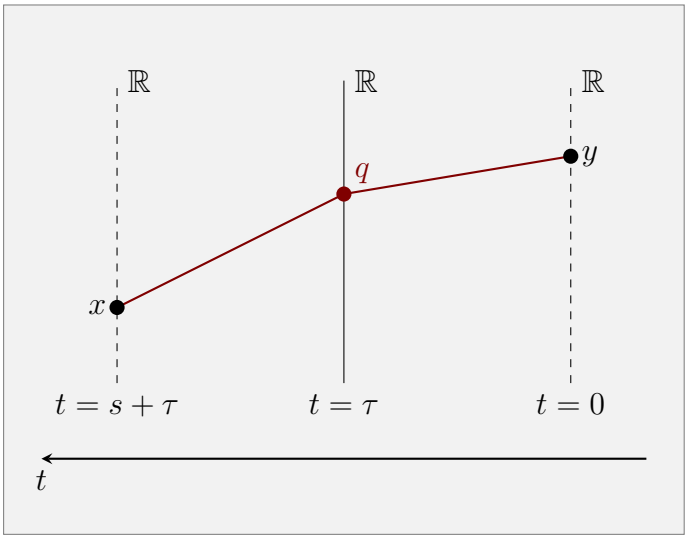
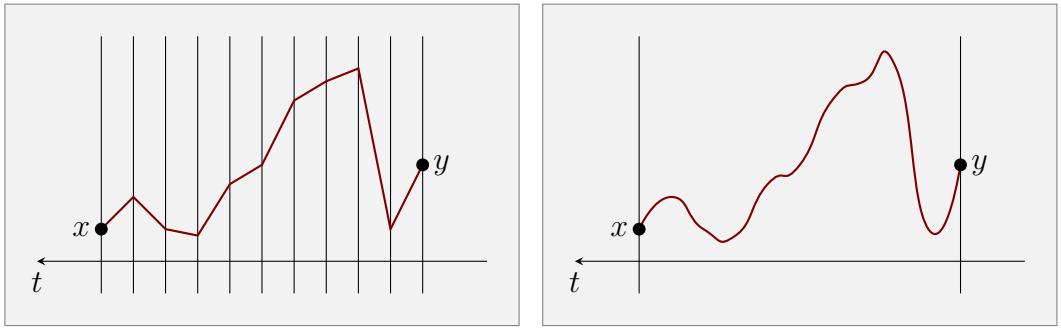


Figure 6.1.: Illustration of propagator composition in $\mathbb{R}$.



(a) **(b)**

Figure 6.2.: Example path in the N slicing case (a) and the continuum limit (b).

The integration can be interpreted as summation over all possibilities. Namely, the particle takes each path (here ignoring their form for intermediate times) from y to each point q and then from the respecting point to x. Each point q has an associated probability density, contributing to the probability amplitude $\langle x|U_{s+\tau}|y\rangle$. Taking two time slicings, the summation/integration runs over all probabilities, that the particle moves like $y \to q_1 \to q_2 \to x$.

The same holds for N time slicings. In the limit $N \to \infty$, the points q_j become smooth paths $q(t)$, the particle travels along, as illustrated in figure 6.2. Each path has a contribution to $\langle x|U_t|y\rangle$, that can be interpreted as proportional to the probability of

the particle taking that path. The contribution is determined by $e^{\frac{i}{\hbar}S[q(t)]}$, where S is the action, arising in the continuum limit.

6.1.3. Gaussian integral and zeta function regularization

It might help to recall some properties of Gaussian integrals in $\mathbb{R}^n$ as well as analytic continuation in complex analysis:

Gaussian integrals in $\mathbb{R}^n$

Lemma 6.1.3.
Let $a, b, c \in \mathbb{R}$ with $a > 0$, then the following equations (each of which is a generalization of the previous one) hold:

$$1)\ \int_{\infty}^{\infty} e^{-x^2}\ dx = \sqrt{\pi} \qquad 2)\ \int_{\infty}^{\infty} e^{-a(x+b)^2}\ dx = \sqrt{\frac{\pi}{a}}$$

$$3)\ \int_{\infty}^{\infty} e^{-ax^2+bx+c}\ dx = \sqrt{\frac{\pi}{a}}\ e^{\frac{b^2}{4a}+c}\ .$$

Proof 6.1.4.
The equations will be proven consecutively to benefit from previous calculations.

1) Using $d(r^2) = 2rdr$ the standard trick is:

$$\begin{aligned}\left(\int_{\infty}^{\infty} e^{-x^2}\ dx\right)^2 &= \int_{\mathbb{R}^2} e^{-(x^2+y^2)}\ dx\ dy = \int_0^{\infty}\int_0^{2\pi} e^{-r^2}\ rd\varphi\ dr \\ &= 2\pi\int_0^{\infty} e^{-r^2}\ rdr = \pi\int_0^{\infty} e^{-r^2}\ dr^2 = \pi\ .\end{aligned}$$

Taking the square root proves the first equation.

2)

$$\begin{aligned}\int_{\infty}^{\infty} e^{-a(x+b)^2}\ dx &= \int_{\infty}^{\infty} e^{-a(x+b)^2}\ d(x+b) = \int_{\infty}^{\infty} e^{-ay^2}\ dy \\ &= \int_{\infty}^{\infty} e^{-a(\frac{z}{\sqrt{a}})^2}\ d\left(\tfrac{z}{\sqrt{a}}\right) = \frac{1}{\sqrt{a}}\int_{\infty}^{\infty} e^{-z^2}\ dz \\ &= \sqrt{\frac{\pi}{a}}\ .\end{aligned}$$

3) Using the previous equations, it is enough to bring the exponent in a suitable form:

$$-ax^2+bx+c = -\left(\sqrt{a}x + \frac{b}{2\sqrt{a}}\right)^2 + \frac{b^2}{4a} + c = -a\left(x+\frac{b}{2a}\right)^2 + \frac{b^2}{4a} + c$$

$$\Rightarrow \qquad e^{-ax^2+bx+c} = e^{-a\left(x+\frac{b}{2a}\right)^2} e^{\frac{b^2}{4a}+c}\ .$$

The rest follows from 2) since the last multiplicand is constant.

□

Corollary 6.1.5.
Let $a, c \in \mathbb{R}$ with $a > 0$ and $b \in \mathbb{R}^n$, then

$$\int_{\mathbb{R}^n} e^{-ax^2+b\cdot x+c}\, dx^n = \left(\frac{\pi}{a}\right)^{\frac{n}{2}}\ e^{\frac{b^2}{4a}+c}\ .$$

Proof 6.1.6.
It is enough to see, that the exponential function factorizes:

$$e^{-ax^2+b\cdot x+c} = e^{\sum_{j=1}^n -ax_j^2+b_jx_j+c} = e^c \prod_{j=1}^n e^{-ax_j^2+b_jx_j}$$

$$\begin{aligned}\Rightarrow \qquad \int_{\mathbb{R}^n} e^{-ax^2+b\cdot x+c}\, dx^n &= e^c \left(\int_\infty^\infty e^{-ax_j^2+b_jx_j}\, dx_j\right)^n \\ &= e^c \prod_{j=1}^n \sqrt{\frac{\pi}{a}} e^{\frac{b_j^2}{4a}} = \left(\frac{\pi}{a}\right)^{\frac{n}{2}}\ e^{\frac{b^2}{4a}+c}\ .\end{aligned}$$

□

Lemma 6.1.7.
Let $A\colon \mathbb{R}^n \to \mathbb{R}^n$ be a symmetric positive definite matrix, then:

$$\int_{\mathbb{R}^n} dx^n\ e^{-a(x,Ax)} = \left(\frac{\pi}{a}\right)^{\frac{n}{2}} \frac{1}{\det(A)^{\frac{1}{2}}}\ .$$

Proof 6.1.8.
Since A is symmetric, there is an orthogonal matrix S, such that $D_A = SAS^T$, where D_A is the diagonalized matrix of A. The property of positive definiteness means, that all values of D_A are positive. To evaluate the integral properly, we understand $\mathbb{R}^n$ as manifold with coordinate map y:

$$\int_{\mathbb{R}^n} dx^n\ e^{-a(x,Ax)} = \int_{\mathbb{R}^n}\ e^{-a(y,Ay)}\ dy^1 \wedge \ldots \wedge dy^n\ .$$

Define the map $f\colon \mathbb{R}^n \to \mathbb{R}^n$ in coordinates by $y(f(p)) = S^T y(p)$. The push forward then is $f_* = S^T$. Since S is orthogonal, i.e. $\det(S^T) = 1$, the pullback of the differential form is:

$$f^*(dy^1 \wedge \ldots \wedge dy^n) = \det(f_*)\ dy^1 \wedge \ldots \wedge dy^n = \det(S^T)\ dy^1 \wedge \ldots \wedge dy^n = dy^1 \wedge \ldots \wedge dy^n$$

Hence the integral becomes ($f^{-1}(\mathbb{R}^n) = \mathbb{R}^n$):

$$\int_{\mathbb{R}^n} dx^n\ e^{-a(x,Ax)} = \int_{\mathbb{R}^n}\ e^{-a(y,Ay)}\ dy^1 \wedge \ldots \wedge dy^n$$

$$
\begin{aligned}
&= \int_{f(f^{-1}(\mathbb{R}^n))} e^{-a(y,Ay)}\ dy^1 \wedge \ldots \wedge dy^n \\
&= \int_{\mathbb{R}^n} f^* \left(e^{-a(y,Ay)}\ dy^1 \wedge \ldots \wedge dy^n \right) \\
&= \int_{\mathbb{R}^n} e^{-a(S^T y, AS^T y)}\ dy^1 \wedge \ldots \wedge dy^n \\
&= \int_{\mathbb{R}^n} dx^n\ e^{-a(S^T x, AS^T x)} = \int_{\mathbb{R}^n} dx^n\ e^{-a(x, SAS^T x)} \\
&= \int_{\mathbb{R}^n} dx^n\ e^{-a(x, D_A x)} = \int_{\mathbb{R}^n} dx^n\ e^{-a\sum_i x_i^2 (D_A)_{ii}} \\
&= \left(\frac{\pi}{a \cdot (D_A)_{11}} \right)^{\frac{1}{2}} \cdot \ldots \cdot \left(\frac{\pi}{a \cdot (D_A)_{nn}} \right)^{\frac{1}{2}} \\
&= \left(\frac{\pi}{a \cdot (D_A)_{11}} \right)^{\frac{n}{2}} \frac{1}{((D_A)_{11} \cdot \ldots \cdot (D_A)_{nn})^{\frac{1}{2}}} \\
&= \left(\frac{\pi}{a} \right)^{\frac{n}{2}} \frac{1}{\det(D_A)^{\frac{1}{2}}} = \left(\frac{\pi}{a} \right)^{\frac{n}{2}} \frac{1}{\det(A)^{\frac{1}{2}}}\ .
\end{aligned}
$$

□

Analytic continuation

First, we will repeat some properties of analytic functions without proof, since they are covered in most courses on complex analysis, e.g. [Swe15].

Let $f\colon U \subset \mathbb{C} \to \mathbb{C}$ be a function. This function is called analytical in $z_0 \in U$, if there is an $r > 0$, such that f can be represented as power series with radius of convergence larger than r:

$$f(z) = \sum_{n \in \mathbb{N}_0} a_n (z - z_0)^n \qquad \forall\ z \in B_r(z_0)\ .$$

While not true in the real case, the complex case allows, that a holomorphic function on an open set $U \subset \mathbb{C}$ is analytic with the taylor series as power series. In fact:

$$f \text{ is analytical on } U \quad \Leftrightarrow \quad f \text{ is holomorphic on } U\ .$$

Theorem 6.1.9 (analytic continuation).
Let $U_1, U_2 \subset \mathbb{C}$ be open and connected and $f_i\colon U_i \to \mathbb{C}$ be holomorphic functions. If there is a $z_0 \in U_1 \cap U_2$ and an $r > 0$, such that

$$f_1(z) = f_2(z) \quad \forall\ z \in B_r(z_0)\ ,$$

then, for every open, connected set $A \subset U_1 \cap U_2$ it holds that

$$f_1(z) = f_2(z) \quad \forall\ z \in A\ .$$

A consequence of this theorem is a unique continuation the the following sense. Let $f\colon U \to \mathbb{R}$ and $F\colon V \to \mathbb{C}$ be holomorphic functions with nonempty open sets $U \subset V \subset \mathbb{C}$, such that

$$f(z) = F(z) \quad \forall\ z \in U\ .$$

The function F is called analytic continuation of f. Assuming there is another function $G\colon V \to \mathbb{C}$, which is also an analytic continuation, then, if V is connected, it holds that

$$F \equiv G \ ,$$

i.e. the analytic continuation is unique. This follows immediately from the last theorem, as $U \subset V$ is open and thus allows for $r > 0$ such that $G(z) = f(z) = F(z)$ on $B_r(z_0)$ for any $z_0 \in U$. Also, since $V \cap V = V$ is open and connected, $F(z) = G(z)$ on V.

The uniqueness of analytic continuation allows to define a complicated function on a small region, where it can be defined easily, and then to continue it to larger domains.

Remark 6.1.10 (Not true in $\mathbb{R}^n$)**.**
The analytic continuation relies on the strong concept of complex differentiation. One might think, that the theorem is trivial in the sense, that it is also true for real functions on $\mathbb{R}^n$. However this is not the case. Recalling the bump functions from theorem F.1.8, we know that j_1 and j_2 are defined on the whole $\mathbb{R}^n$ with $\operatorname{supp}(j_1) = \overline{B_1(0)} \subsetneq \operatorname{supp}(j_2) = \overline{B_2(0)}$. So clearly $j_1 \neq j_2$. Yet on $B_1(y)$ for $\|y\| > 3$ it holds that $j_1(x) = j_2(x) = 0$ for all $x \in B_1(y)$.

Zeta function regularization

From the proof of lemma 6.1.7, it could be hoped, that in taking the limit, creating a countable infinity, the lemma generalizes to

$$\int D[x(t)]\ e^{-\pi(x(t),Ax(t))} \sim \frac{1}{\sqrt{\prod_{j=1}^{\infty} \lambda_j}} \ ,$$

for an operator A with eigen values λ_j. However, in the generic case, the product sequence need not converge. One way to deal with these infinities is the so called zeta function regularization. Additionally to [Zir14] we will also follow [Dun09].

Definition 6.1.11.
Let A be a positive self adjoint endomorphism on the vector space of paths with Hilbert eigen basis and eigen values λ_n , then the **zeta function for A** is defined by

$$\zeta_A(s) = \sum_{n\in\mathbb{N}} \lambda_n^{-s} \qquad \text{for } s \in \mathbb{C} \ .$$

To demand for an Hilbert eigen basis is a technicality that allows to calculate the functional trace as sum of eigen values. From a physical point of view, if the path integral exists, it comes from a limit from the finite dimensional case, where a positive self adjoint operator is diagonizable.

Ignoring the convergence domain for a moment, we can calculate formally:

$$\zeta_A'(s) = -\sum_{n\in\mathbb{N}} \ln(\lambda_n) \cdot \lambda_n^{-s} \quad \Rightarrow \quad \zeta_A'(0) = -\sum_{n\in\mathbb{N}} \ln(\lambda_n) = -\ln\left(\prod_{n\in\mathbb{N}} \lambda_n\right)$$

$$\Rightarrow \qquad \det(A) = \prod_{n\in\mathbb{N}} \lambda_n = e^{-\zeta_A'(0)} \ .$$

Similarly for a quotient of determinants, we obtain formally:

$$\begin{aligned}\zeta_A'(0) - \zeta_B'(0) &= -\sum_{n\in\mathbb{N}} (\ln(\lambda_n) - \ln(\eta_n)) = -\ln\left(\prod_{n\in\mathbb{N}} \frac{\lambda_n}{\eta_n}\right) \\ &= -\ln\left(\frac{\prod_{n\in\mathbb{N}} \lambda_n}{\prod_{n\in\mathbb{N}} \eta_n}\right) \\ \Rightarrow \qquad \frac{\det(A)}{\det(B)} &= \frac{\prod_{n\in\mathbb{N}} \lambda_n}{\prod_{n\in\mathbb{N}} \eta_n} = e^{-(\zeta_A'(0)-\zeta_B'(0))} \ .\end{aligned}$$

Returning to the problem at hand, the whole regularization is necessary because the determinant is not finite. Also, the power series representation is not necessary valid for $s = 0$, as is the case for $\lambda_n = n$. However, sometimes the complex zeta function can be analytically continued, such that $\zeta_A(0)$ and $\zeta_A'(0)$ exist. One famous example is the Riemann zeta function $\lambda_n = n$, for which it has been shown, that $\zeta(0) = -\frac{1}{2}$. If the zeta function (or the difference of two zeta functions) is analytically continuable, we hence define:

$$\det(A) := e^{-\zeta_A'(0)} \qquad \text{and} \qquad \frac{\det(A)}{\det(B)} = e^{-(\zeta_A'(0)-\zeta_B'(0))} \ .$$

This definition enables us to give the determinant a meaning for a larger class of operators, even if the power series diverges.

First application

A first application for the functional Gaussian integral is the first order approximation of a path integral. Recalling the Taylor polynomial for regular functions[2] we want to approximate the action functional. Let S be an action and $x(t)$ be a critical path, i.e. $\frac{\delta S}{\delta x(t)} = 0$. Writing an arbitrary path as $q(t) = x(t) + h(t)$, where $h(t)$ is a small fluctuation around $x(t)$, we approximate:

$$\begin{aligned}S[x(t) + h(t)] &\approx S[x(t)] + \delta S[x(t), h(t)] + \frac{1}{2}\delta^2 S[x(t), h(t)] \\ &= S[x(t)] + \frac{1}{2}\delta^2 S[x(t), h(t)] \ .\end{aligned}$$

The first variation vanishes, since $x(t)$ is a critical path. Rewriting in the standard physical notation, the approximation reads:

$$S[x(t) + h(t)] \approx S[x(t)] + \frac{1}{2}\int dt \int d\tau \ h(t) \frac{\delta^2 S}{\delta x(t)\delta x(\tau)} h(\tau) \ .$$

With the Gaussian integral in mind, we set

$$(h(t), Ah(t)) = \frac{1}{2}\int dt \int d\tau \ h(t) \frac{\delta^2 S}{\delta x(t)\delta x(\tau)} h(\tau)$$

[2] $f(x) \approx f(x_0) + f'(x_0)(x - x_0) + f''(x_0)(x - x_0)^2$ leads to $f(x_0 + \varepsilon) \approx f(x_0) + f'(x_0)\varepsilon + f''(x_0)\varepsilon^2$.

and plug in into the path integral:

$$\begin{aligned}\int D[q(t)]e^{\frac{i}{\hbar}S[q(t)]} &\approx \int D[h(t)]e^{\frac{i}{\hbar}(S[x(t)]+(h(t),Ah(t))\ )}\\ &= e^{\frac{i}{\hbar}S[x(t)]}\int D[h(t)]e^{\frac{i}{\hbar}(h(t),Ah(t))}\\ &= e^{\frac{i}{\hbar}S[x(t)]}\int D[h(t)]e^{\frac{i}{\hbar}\frac{\hbar}{-i\pi}(h(t),\frac{-i\pi}{\hbar}Ah(t))}\\ &= e^{\frac{i}{\hbar}S[x(t)]}\frac{1}{\sqrt{\det(\pi A/i\hbar)}}\ .\end{aligned}$$

6.2. Functional integral for fermions

This section is greatly inspired by [Zir14, V Functional integrals for fermion and bosons] and [AS10, 4 Functional field integral].

6.2.1. Berezin integral in finite dimensions

Let $\bigwedge(V^*)$ be the exterior algebra of the n-dimensional $\mathbb{C}$ dual vector space V^* with dual basis ξ^μ. The generators ξ^μ are called **Grassmann variables** in the physical literature. By lemma D.2.10 we know that elements of the exterior algebra are polynomials of the form:

$$f = f^{(0)} + \sum_\mu f^{(1)}_\mu \xi^\mu + \sum_{\mu<\nu} f^{(2)}_{\mu\nu}\xi^\mu\xi^\nu + \ldots + f^{(n)}_{\mu_1\ldots\mu_n}\xi^{\mu_1}\ldots\xi^{\mu_n}\ ,$$

where we have omitted the wedge product, as is common in the literature.

Definition 6.2.1.
The **partial Grassmann derivative** with respect to a generator ξ^μ is defined linearly by

$$\frac{\partial}{\partial\xi^\mu}1 = 0 \qquad \text{and} \qquad \frac{\partial}{\partial\xi^\mu}\xi^\nu = \delta^\nu_\mu.$$

Lemma 6.2.2.
Let $\{\xi_\mu\}$ be the basis to which $\{\xi^\mu\}$ is dual, then

$$\frac{\partial}{\partial\xi^\mu} = \xi_\mu\lrcorner\ ,$$

where $\xi_\mu\lrcorner$ denotes the interior product.

Proof 6.2.3.
The proof is immediate. □

Corollary 6.2.4.
The partial Grassmann derivative is an anti derivative, i.e.

$$\frac{\partial}{\partial\xi^\mu}(fg) = \left(\frac{\partial}{\partial\xi^\mu}f\right)g + (-1)^{\deg(f)}f\left(\frac{\partial}{\partial\xi^\mu}g\right) \ .$$

Proof 6.2.5.
This is a property of the interior product. □

Corollary 6.2.6.
The partial Grassmann derivatives anti commute:

$$\frac{\partial}{\partial\xi^\mu}\frac{\partial}{\partial\xi^\nu} = -\frac{\partial}{\partial\xi^\nu}\frac{\partial}{\partial\xi^\mu} \qquad \rightarrow \qquad \frac{\partial^2}{\partial\xi^\mu\partial\xi^\mu} = 0 \ .$$

Proof 6.2.7.
This is a property of the interior product. □

Lemma 6.2.8.
Let $\xi^\mu = \sum_\nu A^\mu{}_\nu \zeta^\nu$, then the partial Grassmann derivative transform as the usual partial derivative:

$$\frac{\partial}{\partial\xi^\mu} = \sum_\nu (A^{-1})^\nu{}_\mu \frac{\partial}{\partial\zeta^\nu} \ .$$

Proof 6.2.9.
For the bidual basis it holds that $\xi_\mu = \sum_\nu (A^{-1})^\nu{}_\mu \zeta_\nu$. Plugging that means:

$$\frac{\partial}{\partial\xi^\mu} = \xi_\mu \lrcorner = \sum_\nu (A^{-1})^\nu{}_\mu (\zeta_\nu \lrcorner) = \sum_\nu (A^{-1})^\nu{}_\mu \frac{\partial}{\partial\zeta^\nu} \ .$$

□

So far, the partial Grassmann derivative behaves like a derivative, as the name suggests. Apart from differentiation, the partial Grassmann derivatives can be used to integrate, in a broader sense, on the exterior algebra.

Definition 6.2.10.
Let $f \in \bigwedge(V^*)$ be a polynomial in the exterior algebra, then the **Berezin**

integral $\int_\wedge f\ d\xi^n$ is defined by

$$\int_\wedge f\ d\xi^n \equiv \frac{\partial^n}{\partial\xi^n \dots \partial\xi^1}\ f\ .$$

Remark 6.2.11.
There are different notations for Berezin integrals. Among the most common ones is $\int_F f$. The defining operator $\frac{\partial^n}{\partial\xi^n \dots \partial\xi^1}$ is called **integration form**. Its ordering determines the Berezin integral.

For the Berezin integral there is a more general transformation theorem than lemma 6.2.8:

Theorem 6.2.12 (Berezin change of odd variables)**.**
Let $\varphi\colon \bigwedge(V^*) \to \bigwedge(V^*)$ be an isomorphism and $\xi^\mu = \varphi^\mu(\zeta)$.[3] *Then the formula for the change of variables is:*

$$\int_\wedge f(\xi)\ d\xi^n = \int_\wedge f(\varphi(\zeta))\ \det{}^{-1}\left(\frac{\partial}{\partial\zeta^\mu}\varphi^\nu(\zeta)\right)\ d\zeta^n\ .$$

Since the base space in an algebra, there is a well defined product, which allows to define the exponential function. This also holds true for $\bigwedge(V \oplus V^*)$. Thus we can define:

$$\exp(p) = \sum_n \frac{p^n}{n!}\ , \quad \text{for} \quad p \in \bigwedge(V \oplus V^*)\ .$$

Since the dimension of V is finite, the sum terminates after finitely many terms.

Lemma 6.2.13.
Let $A^\mu{}_\nu$ be the coefficients of $A \in End(V)$ for the basis $\{\xi_\mu\}$ of V and let $\{\zeta^\nu\}$ be a basis of V^*, then:

$$\int_\wedge \exp\left(A^\mu{}_\nu \zeta^\nu \xi_\mu\right)\ d(\xi,\zeta)^n = \det(A)\ ,$$

where the integration form is $\prod_{\mu=1}^n \frac{\partial^2}{\partial\xi_\mu \partial\zeta^\mu}$.

Proof 6.2.14.
Let n be the dimension of V. By the definition of Berezin's integral, only the terms

[3] The index is to be understood as reference to the basis element of $\bigwedge(V^*)$ regarded as vector space. This includes $\varphi^{\mu_1 \cdots \mu_k}(\zeta)$ for $\mu_1 < \ldots < \mu_k$.

of length $2n$, i.e. terms of power n, where all generators appear, do not yield zero:

$$\begin{aligned}
&\int_\wedge \exp(A^\mu{}_\nu \zeta^\nu \xi_\mu)\ d(\xi,\zeta)^n \\
&\quad = \frac{1}{n!}\int_\wedge (A^\mu{}_\nu \zeta^\nu \xi_\mu)^n\ d(\xi,\zeta)^n \\
&\quad = \frac{1}{n!}\int_\wedge \sum_{\pi,\pi'\in\Sigma_n} A^{\pi(1)}{}_{\pi'(1)} \dots A^{\pi(n)}{}_{\pi'(n)} \zeta^{\pi'(1)}\xi_{\pi(1)} \dots \zeta^{\pi'(1)}\xi_{\pi(n)}\ d(\xi,\zeta)^n
\end{aligned}$$

Since the Berezin integral is actually a derivative, that yields 1 for properly ordered generators, we use the anti commutation relations to find:

$$\begin{aligned}
\int_\wedge \exp(A^\mu{}_\nu \zeta^\nu \xi_\mu)\ d(\xi,\zeta)^n &= \frac{1}{n!}\sum_{\pi,\pi'\in\Sigma_n} \mathrm{sgn}(\pi)\mathrm{sgn}(\pi') A^{\pi(1)}{}_{\pi'(1)} \dots A^{\pi(n)}{}_{\pi'(n)} \\
&= \sum_{\pi\in\Sigma_n} \mathrm{sgn}(\pi) A^{\pi(1)}{}_1 \dots A^{\pi(n)}{}_n = \det(A)\ .
\end{aligned}$$

Here, the reordering yields to $\mathrm{sgn}(\tau)$, where $\tau \in \Sigma_{2n}$. Yet the splitting $V \oplus V^*$ is kept, such that τ can be written as product $\tau = \pi \circ \pi'$ for $\pi, \pi' \in \Sigma_n$. □

Corollary 6.2.15.
For later use, we also show that:

$$\int_\wedge \exp\left(-A^\mu{}_\nu \zeta^\nu \xi_\mu\right)\ d(\zeta,\xi)^n = \det(A)\ .$$

Proof 6.2.16.

$$\begin{aligned}
\int_\wedge \exp\left(-A^\mu{}_\nu \zeta^\nu \xi_\mu\right)\ d(\zeta,\xi)^n &= \prod_\mu \frac{\partial^2}{\partial\zeta_\mu \partial\xi_\mu} \left(-A^\mu{}_\nu \zeta^\nu \xi_\mu\right)^n \\
&= (-1)^n \prod_\mu \frac{\partial^2}{\partial\zeta_\mu \partial\xi_\mu} \left(A^\mu{}_\nu \zeta^\nu \xi_\mu\right)^n \\
&= (-1)^n(-1)^n \prod_\mu \frac{\partial^2}{\partial\xi_\mu \partial\zeta_\mu} \left(A^\mu{}_\nu \zeta^\nu \xi_\mu\right)^n \\
&= \int_\wedge \exp\left(A^\mu{}_\nu \zeta^\nu \xi_\mu\right)\ d(\xi,\zeta)^n = \det(A)\ .
\end{aligned}$$

□

This lemma allows to define the Determinant of a matrix. In a different context, the Gaussian Berezin integral can also be used to define another matrix invariant. To be more precise, an invariant of a skew symmetric bilinear form, that can be identified with a matrix.

Definition 6.2.17.
Let $A\colon V \otimes V \to \mathbb{C}$ be a skew symmetric bilinear form. Choosing a dual basis $\{\xi^\mu\}$ of V^*, the bilinear form can be written as $A_{\mu\nu}\xi^\mu \otimes \xi^\nu$, with $A_{\mu\nu} = -A_{\mu\nu}$. Then, the **Pfaffian** of A is defined by

$$\mathrm{Pf}(A) = \int_\wedge \exp(A_{\mu\nu}\xi^\mu\xi^\nu)\, d\xi^n \ ,$$

with the standard integration form $\frac{\partial^n}{\partial\xi^n \ldots \partial\xi^1}$.

Corollary 6.2.18.
If $\dim(V) = n$ *is odd, then* $\mathrm{Pf}(A) = 0$.

Proof 6.2.19.
The terms of $\exp(A_{\mu\nu}\xi^\mu\xi^\nu)$ that are non-zero under Berezin integration have to have length n. Yet since $(A_{\mu\nu}\xi^\mu\xi^\nu)^k$ yields only terms of even length but n is odd, the only possible result of Berezin integration is 0. □

The Pfaffian is closely related to the determinant. Let $I\colon V^* \to V$ be the isomorphism induced by the choice of bases (in the case of euclidean vector spaces it becomes canonical). Then the associated matrix of $A = A_{\mu\nu}\xi^\mu \otimes \xi^\nu$ is

$$\widetilde{A} = (I \otimes \mathbb{1})(A) = A_{\mu\nu} I(\xi^\mu) \otimes \xi^\nu = \sum_{\mu,\nu} A_{\mu\nu}\xi_\mu \otimes \xi^\nu \equiv \widetilde{A}^\mu{}_\nu \xi_\mu \otimes \xi^\nu = \widetilde{A} \ .$$

Theorem 6.2.20 (Cayley's theorem)**.**
If $\dim(V) = n$ *is even, then* $\mathrm{Pf}(A)^2 = \det(\widetilde{A})$.

6.2.2. Fermionic coherent states

While coherent states in the bosonic case are discussed quite early in quantum mechanics (in terms of excitations), cohernet states of fermions pose much more difficulty. In the many particle case, a fermionic coherent state would need to fulfill the relation $c_j|\xi\rangle = \xi_i$. However, because of the anti commuting annihilation operators, ξ_i cannot be a number:

$$c_i c_j|\xi\rangle = \xi_j c_i|\xi\rangle = \xi_j\xi_i|\xi\rangle \overset{!}{=} -c_j c_i|\xi\rangle = \xi_i\xi_j|\xi\rangle \quad \Rightarrow \quad \xi_i\xi_j \overset{?}{=} -\xi_j\xi_i.$$

To formulate a theory of fermionic coherent states, we loosen our restrictions from the previous subsection and allow for possibly infinite dimensional spaces V and V^*. We consider creation and annihilation operators $c_i^\dagger, c_j \in \mathrm{Cl}(\mathcal{H} \oplus \mathcal{H}^*)$ and Grassmann generators $\xi_i, \overline{\xi}_j \in \bigwedge(V \oplus V^*)$. The notation for $\overline{\xi}_i \in \mathcal{H}^*$ is motivated by the following definition:

Definition 6.2.21.
The **complex conjugation** in the exterior algebra is defined by $\xi_i \mapsto \overline{\xi}_i$ and $\overline{\xi}_i \mapsto \overline{\overline{\xi}_i} \equiv \xi_i$, such that:

$$\overline{\alpha\xi_i + \beta\xi_j} = \overline{\alpha}\overline{\xi}_i + \overline{\beta}\overline{\xi}_j \qquad \text{and} \qquad \overline{\xi_i\xi_j} = \overline{\xi}_i\overline{\xi}_j \ .$$

Corollary 6.2.22.
It holds that

$$\int_\wedge d(\overline{\xi},\xi)^n \ \exp\left(-\sum_\mu \overline{\xi}_\mu\xi_\mu\right) = 1 \ ,$$

where we have chosen the more conventional integration form $\int_\wedge d(\overline{\xi},\xi)^n = \prod_j \frac{\partial^2}{\partial\overline{\xi}_j\partial\xi_j}$.

Proof 6.2.23.
This is a special case of corollary 6.2.15 using $A^\mu{}_\nu = \delta^\mu_\nu$. □

Furthermore, we make the following convention:

Grassmann generators and creation /annihilation operators anticommute, i.e. $\xi_i c_j^\dagger = -c_j^\dagger\xi_i$ etc. .

Before we continue with the Berezin integral, we consider some important states, called coherent states:

Definition 6.2.24.
The states $|\xi\rangle$ defined by

$$|\xi\rangle = \exp\left(-\sum_\mu \xi_\mu c_\mu^\dagger\right)|0\rangle \equiv \exp(\boldsymbol{\xi})|0\rangle$$

are called **fermionic coherent states**, where we defined $\boldsymbol{\xi} = \sum_\mu c_\mu^\dagger\xi_\mu$. Furthermore we define $\overline{\boldsymbol{\xi}} = \sum_\mu c_\mu\overline{\xi}_\mu$

For the adjoint state we find:

$$\langle\xi| = |\xi\rangle^\dagger = |0\rangle^\dagger\exp(\boldsymbol{\xi})^\dagger = \langle 0|\exp\left(\left(\sum_\mu c_\mu^\dagger\xi_\mu\right)^\dagger\right) = \langle 0|\exp\left(\sum_\mu \overline{\xi}_\mu c_\mu\right)$$
$$= \langle 0|\exp(-\overline{\boldsymbol{\xi}}) \ .$$

Before we show the properties of coherent states, we need a statement about the exponential function.

Corollary 6.2.25.
The following commutator relations hold:

i) $[c_\mu^\dagger, c_\nu^\dagger \xi_\nu] = 0 \ , \quad [c_\mu, c_\nu^\dagger \xi_\nu] = \delta_{\mu\nu}\xi_\nu \ , \quad [c_\mu^\dagger, c_\nu \overline{\xi}_\nu] = \delta_{\mu\nu}\overline{\xi}_\nu .$

ii) $[\xi_\mu, c_\nu^\dagger \xi_\nu] = 0 \ , \quad [\overline{\xi}_\mu, c_\nu^\dagger \xi_\nu] = 0 \ , \quad [\xi_\mu, c_\nu \overline{\xi}_\nu] = 0 \ , \quad [\overline{\xi}_\mu, c_\nu \overline{\xi}_\nu] = 0 \ .$

Proof 6.2.26.

i) The first equation hold because of anti-commutativity and two exchanges. For the second and third commutator, we calculate:

$$[c_\mu, c_\nu^\dagger \xi_\nu] = c_\mu c_\nu^\dagger \xi_\nu - c_\nu^\dagger \xi_\nu c_\mu = (c_\mu c_\nu^\dagger + c_\nu^\dagger c_\mu)\xi_\nu = \delta_{\mu\nu}\xi_\nu \ .$$

$$[c_\mu^\dagger, c_\nu \overline{\xi}_\nu] = c_\mu^\dagger c_\nu \overline{\xi}_\nu - c_\nu \overline{\xi}_\nu c_\mu^\dagger = (c_\mu^\dagger c_\nu + c_\nu c_\mu^\dagger)\overline{\xi}_\nu = \delta_{\mu\nu}\overline{\xi}_\nu$$

ii) All commutators vanish, since there are always two exchanges.

□

Lemma 6.2.27.
It holds that:

$$[\boldsymbol{\xi}, \overline{\boldsymbol{\zeta}}] = \sum_\mu \overline{\zeta}_\mu \xi_\mu \ , \qquad [\boldsymbol{\xi}, \boldsymbol{\zeta}] = 0 = [\overline{\boldsymbol{\xi}}, \overline{\boldsymbol{\zeta}}] \ ,$$

and that higher commutators vanish.

Proof 6.2.28.
We calculate the following commutators, using the anti-commutator relations:

$$\begin{aligned} -[\boldsymbol{\xi}, \overline{\boldsymbol{\zeta}}] = [\boldsymbol{\xi}, -\overline{\boldsymbol{\zeta}}] &= \left[\sum_\mu c_\mu^\dagger \xi_\mu \ , \ \sum_\rho \overline{\zeta}_\rho c_\rho \right] = \sum_{\mu,\rho} c_\mu^\dagger \xi_\mu \overline{\zeta}_\rho c_\rho - \overline{\zeta}_\rho c_\rho c_\mu^\dagger \xi_\mu \\ &= -\sum_{\mu,\rho} \overline{\zeta}_\rho (c_\mu^\dagger c_\rho + c_\rho c_\mu^\dagger)\xi_\mu \\ &= -\sum_{\mu,\rho} \overline{\zeta}_\rho \ \delta_{\mu\rho} \ \xi_\mu = -\sum_\mu \overline{\zeta}_\mu \xi_\mu \ , \end{aligned}$$

For the remaining commutators:

$$[c_\mu^\dagger \xi_\mu, c_\eta^\dagger \zeta_\eta] = c_\mu^\dagger \xi_\mu c_\eta^\dagger \zeta_\eta - c_\eta^\dagger \zeta_\eta c_\mu^\dagger \xi_\mu = c_\mu^\dagger \xi_\mu c_\eta^\dagger \zeta_\eta - c_\mu^\dagger \xi_\mu c_\eta^\dagger \zeta_\eta = 0$$

$$\Rightarrow \qquad [\boldsymbol{\xi}, \boldsymbol{\zeta}] = \left[\sum_\mu c_\mu^\dagger \xi_\mu \ , \ \sum_\eta c_\eta^\dagger \zeta_\eta \right] = 0$$

$$\text{and similarly} \quad [\overline{\boldsymbol{\xi}}, \overline{\boldsymbol{\zeta}}] = \left[\sum_\nu \overline{\xi}_\nu c_\nu \ , \ \sum_\rho \overline{\zeta}_\rho c_\rho \right] = 0 \ .$$

Higher commutators vanish, because of corollary 6.2.25. □

Lemma 6.2.29.
It holds that $\exp\left(\sum_\mu c_\mu^\dagger \xi_\mu\right) = \prod_\mu (1 + c_\mu^\dagger \xi_\mu)$.

Proof 6.2.30.
Because the variable pairs $c_\mu^\dagger \xi_\mu$ commute, we can write

$$\exp\left(\sum_{\mu\neq\nu} c_\mu^\dagger \xi_\mu + c_\nu^\dagger \xi_\nu\right) = \exp\left(\sum_\mu c_\mu^\dagger \xi_\mu\right) \exp\left(c_\nu^\dagger \xi_\nu\right) \ .$$

For the case $n = 1$ it is clear that $\exp\left(c_\mu^\dagger \xi_\mu\right) = (1 + c_\mu^\dagger \xi_\mu)$.[4]Because of the factorization of exponential functions above, the induction is fairly easy:

$$\begin{aligned}\exp\left(\sum_{\mu\neq\nu} c_\mu^\dagger \xi_\mu + c_\nu^\dagger \xi_\nu\right) &= \exp\left(\sum_\mu c_\mu^\dagger \xi_\mu\right) \exp\left(c_\nu^\dagger \xi_\nu\right) \\ &= \prod_{\mu\neq\nu} (1 + c_\mu^\dagger \xi_\mu)(1 + c_\nu^\dagger \xi_\nu) = \prod_\mu (1 + c_\mu^\dagger \xi_\mu) \ .\end{aligned}$$

□

Lemma 6.2.31.
Fermionic coherent states have the following properties:

$$i) \quad c_\mu|\xi\rangle = \xi_\mu|\xi\rangle \qquad ii) \quad c_\nu^\dagger|\xi\rangle = -\frac{\partial}{\partial\xi_\nu}|\xi + c_\nu^\dagger \xi_\nu\rangle$$

$$iii) \quad \langle\zeta|\xi\rangle = \exp\left(\sum_\mu \overline{\zeta}_\mu \xi_\mu\right) \ .$$

Proof 6.2.32.

i) For the first property we observe for a coherent state $|\mu\rangle$ not containing μ:

$$c_\mu(1 - \xi_\mu c_\mu^\dagger)|\mu\rangle = -c_\mu \xi_\mu c_\mu^\dagger|\mu\rangle = \xi_\mu c_\mu c_\mu^\dagger|\mu\rangle = \xi_\mu|\mu\rangle \ ,$$

$$\xi_\mu^2 = 0 \qquad \Rightarrow \qquad \xi_\mu|\mu\rangle = \xi_\mu(1 - \xi_\mu c_\mu^\dagger)|\mu\rangle$$

Using the product representation of the exponential function yields:

$$\begin{aligned}c_\mu|\xi\rangle = c_\mu \exp\left(\sum_\nu c_\nu^\dagger \xi_\nu\right)|0\rangle = \\ c_\mu \prod_\nu (1 - \xi_\nu c_\nu^\dagger)|0\rangle = c_\mu(1 - \xi_\mu c_\mu^\dagger) \prod_{\nu\neq\mu} (1 - \xi_\nu c_\nu^\dagger)|0\rangle\end{aligned}$$

[4]Although the variable pairs $c_\mu^\dagger \xi_\mu$ commute, they still vanish upon quadration. This is also the difference to the exponential function in symmetric variables, where the case for $n = 1$ is different.

$$= \xi_\mu \prod_{\nu\neq\mu}(1-\xi_\nu c_\nu^\dagger)|0\rangle = \xi_\mu(1-\xi_\mu c_\mu^\dagger)\prod_{\nu\neq\mu}(1-\xi_\nu c_\nu^\dagger)|0\rangle$$
$$= \ldots = \xi_\mu|\xi\rangle \ .$$

ii) For the second property we use:

$$-\frac{\partial}{\partial\xi_\nu}(1+c_\nu^\dagger\xi_\nu) = c_\nu^\dagger$$

and the commutativity of variable pairs from corollary 6.2.25:

$$-\frac{\partial}{\partial\xi_\nu}\exp(\xi+c_\nu^\dagger\xi_\nu) = -\frac{\partial}{\partial\xi_\nu}(1+c_\nu^\dagger\xi_\nu)\prod_{\mu\neq\nu}(1+c_\mu^\dagger\xi_\mu)$$
$$= c_\nu^\dagger\prod_{\mu\neq\nu}(1+c_\mu^\dagger\xi_\mu) = c_\nu^\dagger\exp(\xi) \ .$$

iii) The last property is a direct calculation using

$$\langle 0|(1+\overline{\zeta}_\nu c_\nu)(1+c_\mu^\dagger\xi_\mu)|0\rangle = \langle 0|(1+c_\mu^\dagger\xi_\mu+\overline{\zeta}_\nu c_\nu+\overline{\zeta}_\nu c_\nu c_\mu^\dagger\xi_\mu)|0\rangle$$
$$= \langle 0|1+\overline{\zeta}_\nu\delta_{\mu_\nu}\xi_\mu|0\rangle \ .$$

Thus we find:

$$\langle\zeta|\xi\rangle = \left\langle 0\left|\prod_{\mu,\nu}(1+\overline{\zeta}_\nu c_\nu)(1+c_\mu^\dagger\xi_\mu)\right|0\right\rangle = \left\langle 0\left|\prod_{\mu,\nu}1+\overline{\zeta}_\nu\delta_{\mu_\nu}\xi_\mu\right|0\right\rangle$$
$$= \left\langle 0\left|\prod_\mu 1+\overline{\zeta}_\mu\xi_\mu\right|0\right\rangle = \exp\left(\sum_\mu\overline{\zeta}_\mu\xi_\mu\right) \ .$$

□

Theorem 6.2.33.
The identity operator on the Fock space can be expressed as follows:

$$\mathbb{1} = \int_\wedge d(\overline{\xi},\xi)\ \exp\left(-\sum_\mu\overline{\xi}_\mu\xi_\mu\right)|\xi\rangle\langle\xi| \ .$$

Proof 6.2.34.
We consider the Hilbert basis of the Fock space:

$$|n_1,\ldots\rangle = (c_1^\dagger)^{n_1}\ldots(c_k^\dagger)^{n_k}\ldots|0\rangle \ .$$

Then by definition and lemma 6.2.31 the following holds:

$$\langle m_1,\ldots|n_1,\ldots\rangle = \prod_j\delta_{m_jn_j} \ ,$$

$$\langle m_1,\ldots|\xi\rangle = \langle 0|\ldots\xi_2^{m_2}\xi_1^{m_1}|\xi\rangle = \ldots\xi_1^{m_1} \ ,$$

$$\langle \xi | n_1, \ldots \rangle = \langle \xi | \overline{\xi}_1^{n_1} \ldots | 0 \rangle = \overline{\xi}_1^{n_1} \ldots \ .$$

Next we consider the operator $\mathcal{O} = \int_\wedge d(\overline{\xi}, \xi) \ \exp\left(-\sum_\mu \overline{\xi}_\mu \xi_\mu\right) |\xi\rangle\langle\xi|$:

$$\langle m_1, \ldots | \mathcal{O} | \, n_1, \ldots \rangle = \left\langle m_1, \ldots \left| \int_\wedge d(\overline{\xi}, \xi) \ \exp\left(-\sum_\mu \overline{\xi}_\mu \xi_\mu\right) |\xi\rangle\langle\xi| \right| n_1, \ldots \right\rangle \ .$$

Since each term of $\exp\left(-\sum_\mu \overline{\xi}_\mu \xi_\mu\right)$ is even in the total number of generators, it commutes with $\langle m_1, \ldots |$. Furthermore, $(\overline{\xi}_j)^{n_j}$ and $(\xi_j)^{m_j}$ commute with the coherent states, as well as $\exp\left(-\sum_\mu \overline{\xi}_\mu \xi_\mu\right)$, for the same reason. Hence:

$$\begin{aligned}
\langle m_1, \ldots | \mathcal{O} | \, n_1, \ldots \rangle &= \int_\wedge d(\overline{\xi}, \xi) \ \exp\left(-\sum_\mu \overline{\xi}_\mu \xi_\mu\right) \langle m_1, \ldots | \xi \rangle \langle \xi | n_1, \ldots \rangle \\
&= \int_\wedge d(\overline{\xi}, \xi) \ \exp\left(-\sum_\mu \overline{\xi}_\mu \xi_\mu\right) \ldots \xi_1^{m_1} \overline{\xi}_1^{n_1} \ldots \\
&= \int_\wedge d(\overline{\xi}, \xi) \ \exp\left(-\sum_\mu \overline{\xi}_\mu \xi_\mu\right) \xi_1^{m_1} \overline{\xi}_1^{n_1} \xi_2^{m_2} \overline{\xi}_2^{n_2} \ \ldots \ ,
\end{aligned}$$

where we exchanged the order in the last line (always an even number of exchanges).

Furthermore:

$$\begin{aligned}
&\int_\wedge d(\overline{\xi}, \xi)^k \ \exp\left(-\sum_\mu \overline{\xi}_\mu \xi_\mu\right) \xi_1^{m_1} \overline{\xi}_1^{n_1} \xi_2^{m_2} \overline{\xi}_2^{n_2} \ \ldots \\
&\quad = \int_\wedge d(\overline{\xi}, \xi)^k \ \left(-\sum_\mu \overline{\xi}_\mu \xi_\mu\right)^k \xi_1^{m_1} \overline{\xi}_1^{n_1} \xi_2^{m_2} \overline{\xi}_2^{n_2} \ \ldots \\
&\quad = \int_\wedge d(\overline{\xi}, \xi)^k \ \left(-\sum_\mu \overline{\xi}_\mu \xi_\mu\right) (-1)^k \xi_1^{m_1} \overline{\xi}_1^{n_1} \xi_2^{m_2} \overline{\xi}_2^{n_2} \ \ldots \\
&\quad = \int_\wedge d(\overline{\xi}, \xi)^k \ \left(-\sum_\mu \overline{\xi}_\mu \xi_\mu\right) \overline{\xi}_1^{n_1} \xi_1^{m_1} \overline{\xi}_2^{n_2} \xi_2^{m_2} \ \ldots
\end{aligned}$$

$$\Rightarrow \qquad \langle m_1, \ldots | \mathcal{O} | \, n_1, \ldots \rangle = \int_\wedge d(\overline{\xi}, \xi) \ \left(-\sum_\mu \overline{\xi}_\mu \xi_\mu\right) \overline{\xi}_1^{n_1} \xi_1^{m_1} \overline{\xi}_2^{n_2} \xi_2^{m_2} \ \ldots \ .$$

Assuming at least for one j, that $m_j \neq n_j$, since $\exp\left(\sum_\mu \overline{\xi}_\mu \xi_\mu\right)$ consists of generator pairs, there is no term containing all generators (needed for the Berezin integral), since ξ_j has no "partner" $\overline{\xi}_j$ or vice versa. Thus we can assume $m_j = n_j$ for all j. Since all generator pairs commute, the terms of proper length can always be ordered, such that

$$\langle m_1, \ldots | \mathcal{O} | \, n_1, \ldots \rangle = \prod_j \delta_{m_j n_j} = \langle m_1, \ldots | n_1, \ldots \rangle \ .$$

□

Lemma 6.2.35.
Let $|\psi\rangle$ and $|\psi\rangle$ be states, then it holds that:

$$\begin{aligned}
&\int_\wedge d(\overline{\xi},\xi)\ \exp\left(-\sum_\mu \overline{\xi}_\mu\xi_\mu\right)\langle\psi|\xi\rangle\langle\xi|\phi\rangle \\
&\quad=\int_\wedge d(\overline{\xi},\xi)\ \exp\left(-\sum_\mu \overline{\xi}_\mu\xi_\mu\right)\langle-\xi|\phi\rangle\langle\psi|\xi\rangle \\
&\quad=\int_\wedge d(\overline{\xi},\xi)\ \exp\left(-\sum_\mu \overline{\xi}_\mu\xi_\mu\right)\langle\xi|\phi\rangle\langle\psi|-\xi\rangle\ .
\end{aligned}$$

Proof 6.2.36.
Let $|n_1,\ldots\rangle$ denote an arbitrary Fock space Hilbert base state. By linearity[5], we only need to consider $\langle m_1,\ldots|\xi\rangle\langle\xi|n_1,\ldots\rangle$. To keep the notation short, we define $\int \equiv \int_\wedge d(\overline{\xi},\xi)\ \exp\left(-\sum_\mu \overline{\xi}_\mu\xi_\mu\right)$ in this proof. Using the proof of theorem 6.2.33 we find:

$$\begin{aligned}
\int\langle m_1,\ldots|\xi\rangle\langle\xi|n_1,\ldots\rangle &= \int\ldots\xi_1^{m_1}\overline{\xi}_1^{n_1}\ldots = \int\prod_j\delta_{m_jn_j}\ldots\xi_1^{m_1}\overline{\xi}_1^{n_1}\ldots \\
&= \int\ldots\xi_1^{n_1}\overline{\xi}_1^{n_1}\ldots\ .
\end{aligned}$$

If $\overline{\xi}_1^{n_1}\ldots$ is of even length, so is $\ldots\xi_1^{n_1}$, such that $\overline{\xi}_1^{n_1}\ldots = (-\overline{\xi})_1^{n_1}\ldots$ and thus:

$$\ldots\xi_1^{n_1}\overline{\xi}_1^{n_1}\ldots = \overline{\xi}_1^{n_1}\ldots\ldots\xi_1^{n_1} = (-\overline{\xi})_1^{n_1}\ldots\ldots\xi_1^{n_1}\ .$$

Similarly for odd length, $\overline{\xi}_1^{n_1}\ldots = -((-\overline{\xi})_1^{n_1}\ldots)$ and:

$$\ldots\xi_1^{n_1}\overline{\xi}_1^{n_1}\ldots = -(\overline{\xi}_1^{n_1}\ldots\ldots\xi_1^{n_1}) = (-\overline{\xi})_1^{n_1}\ldots\ldots\xi_1^{n_1}\ .$$

So in general we see that:

$$\begin{aligned}
\int\langle m_1,\ldots|\xi\rangle\langle\xi|n_1,\ldots\rangle &= \int\ldots\xi_1^{n_1}\overline{\xi}_1^{n_1}\ldots = \int(-\overline{\xi})_1^{n_1}\ldots\ldots\xi_1^{n_1} \\
&= \ldots = \int\langle-\xi|n_1,\ldots\rangle\langle m_1,\ldots|\xi\rangle\ .
\end{aligned}$$

The second equality can be shown in the same way. □

6.2.3. Functional integral for fermions: The grand canonical partition function

To illustrate the functional integral for fermions, we consider the grand canonical partition function. As a reminder, in classical statistical mechanics, the grand canonical

[5] assumed to exchange with Berezin integration by continuity. Though this lemma is used for the trace only, where one could defined the trace naively by allowing only the occupation number basis, the continuity still would be needed for the calculations.

partition function is given by

$$\mathcal{Z} = \sum_r e^{-\beta(E_r - \mu N_r)} \ , \quad \text{with } \beta = k_B T \ .$$

In elementary quantum mechanics, we would define an extended Hamilton operator H with normalized eigen states $|r\rangle$, such that $H|r\rangle = (E_r - \mu N_r)$, and find that:

$$\begin{aligned}\mathcal{Z} &= \sum_r e^{-\beta(E_r-\mu N_r)} = \sum_r e^{-\beta(E_r-\mu N_r)}\langle r|r\rangle = \sum_r \langle r|e^{-\beta(E_r-\mu N_r)}|r\rangle \\ &= \sum_r \langle r|e^{-\beta H}|r\rangle = \mathrm{Tr}(e^{-\beta H}) \ .\end{aligned}$$

This is taken as general definition of the partition function, such that we want to develop a functional integral representation for $\mathcal{Z} = \mathrm{Tr}(e^{-\beta H})$. We assume, that H is even in the number of creation/annihilation operators[6], and that it is already normal ordered, otherwise we could consider $:H:$ instead. In case of a fermionic system, with an arbitrary ordered Hilbert basis $|n\rangle$, using theorem 6.2.33 and lemma 6.2.35, where $\langle\psi| = \langle n|$ and $|\phi\rangle = e^{-\beta H}|n\rangle$ we calculate:

$$\begin{aligned}\mathcal{Z} &= \mathrm{Tr}(e^{-\beta H}) \\ &= \sum_n \langle n|e^{-\beta H}|n\rangle = \sum_n \int_\wedge d(\overline{\xi},\xi)\ \exp\left(-\sum_\mu \overline{\xi}_\mu \xi_\mu\right) \langle n|\xi\rangle\langle\xi|e^{-\beta H}|n\rangle \\ &= \int_\wedge d(\overline{\xi},\xi)\ \exp\left(-\sum_\mu \overline{\xi}_\mu \xi_\mu\right) \sum_n \langle\xi|e^{-\beta H}|n\rangle\langle n|-\xi\rangle \\ &= \int_\wedge d(\overline{\xi},\xi)\ \exp\left(-\sum_\mu \overline{\xi}_\mu \xi_\mu\right) \langle\xi|e^{-\beta H}|-\xi\rangle \ .\end{aligned}$$

In the last line we used the completeness relation $\sum_n |n\rangle\langle n| = \mathbb{1}$. Having a theory of fields in the continuum limit in mind and recalling the general definition of field operators, being an infinite linear combination of operators, we understand ξ and $\overline{\xi}$ as "vectors", with $\overline{\xi}\xi = \sum_\mu \overline{\xi}_\mu \xi_\mu$. Following the general procedure, we split the exponential and insert completeness relations:

$$\begin{aligned}\mathcal{Z} = \int_\wedge d(\overline{\xi}_1,\xi_1)\ldots\int_\wedge d(\overline{\xi}_M,\xi_M)\ &\exp(-\overline{\xi}_M\xi_M)\ldots\exp(-\overline{\xi}_1\xi_1) \\ &\langle\xi_M|e^{-\frac{1}{M}\beta H}|\xi_{M-1}\rangle\langle\xi_{M-1}|e^{-\frac{1}{M}\beta H}|\xi_{M-2}\rangle\ldots\langle\xi_1|e^{-\frac{1}{M}\beta H}|-\xi_M\rangle \\ = \int_\wedge d(\overline{\xi}_1,\xi_1)\ldots\int_\wedge d(\overline{\xi}_M,\xi_M)\ &\exp\left(-\sum_{j=1}^M \overline{\xi}_j\xi_j\right) \\ &\langle\xi_M|e^{-\frac{1}{M}\beta H}|\xi_{M-1}\rangle\langle\xi_{M-1}|e^{-\frac{1}{M}\beta H}|\xi_{M-2}\rangle\ldots\langle\xi_2|e^{-\frac{1}{M}\beta H}|-\xi_1\rangle \ .\end{aligned}$$

In the second line, we used that pairs of Grassmann generators commute. Next we consider the essential building block $\langle\xi|e^{-\frac{1}{M}\beta H}|\zeta\rangle$. We will expand the exponential up to the first order, as we intend to set $M \to \infty$. Since H is normal ordered, we can exchange creation/annihilation operators with the corresponding Grassmann generators,

[6] Which is true most of the times.

because of lemma 6.2.31. Then, because of evenness w.r.t. the number of generator, we can use commutation with fermionic coherent states:

$$\begin{aligned}\langle\xi|e^{-\frac{1}{M}\beta H}|\zeta\rangle &\approx \langle\xi|1-\tfrac{1}{M}\beta H(c^\dagger,c)|\zeta\rangle = \langle\xi|1-\tfrac{1}{M}\beta H(\overline{\xi},\zeta)|\zeta\rangle \\ &= \langle\xi|\zeta\rangle(1-\tfrac{1}{M}\beta H(\overline{\xi},\zeta)) = e^{\overline{\xi}\zeta}(1-\tfrac{1}{M}\beta H(\overline{\xi},\zeta)) \\ &\approx e^{\overline{\xi}\zeta}e^{-\frac{1}{M}\beta H(\overline{\xi},\zeta)} = e^{\overline{\xi}\zeta-\frac{1}{M}\beta H(\overline{\xi},\zeta)} \ .\end{aligned}$$

Plugging in these results and demanding the condition $-\xi_M = \xi_1$, we find[7]:

$$\begin{aligned}\mathcal{Z} &= \int_\wedge d(\overline{\xi}_1,\xi_1)\dots\int_\wedge d(\overline{\xi}_M,\xi_M)\ \exp\left(\sum_{j=2}^{M}(-\overline{\xi}_j\xi_j+\overline{\xi}_j\xi_{j-1})\ -\frac{1}{M}\beta H(\overline{\xi}_j,\xi_{j-1})\right) \\ &\overset{\delta\equiv\frac{\beta}{M}}{=} \int_\wedge d(\overline{\xi}_1,\xi_1)\dots\int_\wedge d(\overline{\xi}_M,\xi_M)\ \exp\left(-\delta\sum_{j=2}^{M}\overline{\xi}_j\frac{\xi_j-\xi_{j-1}}{\delta}+H(\overline{\xi}_j,\xi_{j-1})\right) \ .\end{aligned}$$

Assuming the continuum limit $M\to\infty$ to exist, which leads to $\delta\to 0$, we understand $\xi_j\to\psi(j)$ and $\overline{\xi}_j\to\overline{\psi}(j)$ as continuous fields. Also:

$$\int_\wedge d(\overline{\xi}_1,\xi_1)\dots\int_\wedge d(\overline{\xi}_M,\xi_M)\longrightarrow\int D(\overline{\psi},\psi)\ ,$$

$$\delta\sum_{j=1}^{M}\longrightarrow\int_0^\beta d\tau \qquad\text{and}\qquad \overline{\xi}_j\frac{\xi_j-\xi_{j-1}}{\delta}\longrightarrow\overline{\psi}\partial_\tau\psi\ .$$

Thus we have found a field integral for the partition function:

$$\boxed{\mathcal{Z}=\int D(\overline{\psi},\psi)\ e^{-\int_0^\beta d\tau\ \overline{\psi}(\tau)\partial_\tau\psi(\tau)+H(\overline{\psi}(\tau),\psi(\tau))}\ .}$$

The condition $-\xi_M=\xi_1$ becomes a boundary condition:

$$-\psi(\beta)=\psi(0)\qquad\text{and}\qquad-\overline{\psi}(\beta)=\overline{\psi}(0)\ .$$

6.3. Functional integral for bosons

As the section 6.2, this is greatly inspired by [Zir14, V Functional integrals for fermion and bosons] and [AS10, 4 Functional field integral].

6.3.1. Coherent states for bosons

While fermionic coherent states needed a new concept of "numbers" to realize the concept of coherent states, bosonic coherent states arise more natural.

Definition 6.3.1.

[7] Because of the commutativity of Grassmann pairs, we can use the usual power law for the exponential functions.

Let $\phi_j \in \mathbb{C}$. The **bosonic coherent state** $|\phi\rangle$ is defined by

$$|\phi\rangle = \exp\left(\sum_j \phi_j a_j^\dagger\right)|0\rangle \ .$$

As the name suggests, bosonic coherent states have the same properties as their fermionic counterparts. However, since the bosonic theory is generally symmetric, while the fermionic one is generally antisymmetric, the details differ. For example, adopting the previous convention $\boldsymbol{\phi} \equiv \sum_j \phi_j a_j^\dagger$, the bra-state becomes:

$$\langle\phi| = |\phi\rangle^\dagger = \langle 0|\exp(\boldsymbol{\phi}^\dagger) = \langle 0|\exp\left(\sum_j \overline{\phi}_j a_j\right) \ .$$

Lemma 6.3.2.
Bosonic coherent states have the following properties:

$$i) \quad a_\mu|\phi\rangle = \phi_\mu|\phi\rangle \qquad ii) \quad a_\nu^\dagger|\phi\rangle = \frac{\partial}{\partial\phi_\nu}|\phi + a_\nu^\dagger\phi_\nu\rangle$$

$$iii) \quad \langle\theta|\phi\rangle = \exp\left(\sum_\mu \overline{\theta}_\mu\phi_\mu\right) \ .$$

Proof 6.3.3.

i) For the first property, we observe. that a_μ and ϕ_ν commute in general, and a_μ and $a_\nu^\dagger$ if $\mu \neq \nu$. Since the $a_\nu^\dagger$ commute with each other, we can order $\exp(\boldsymbol{\phi})$. For that reason, if $\boldsymbol{\phi}' = \sum_{\nu\neq\mu}\phi_\nu a_\nu^\dagger$, i.e. $\boldsymbol{\phi} = \phi_\mu a_\mu^\dagger + \boldsymbol{\phi}'$, it holds that $\exp(\boldsymbol{\phi}) = \exp(\boldsymbol{\phi}')\exp(\phi_\mu a_\mu^\dagger)$. Also because of $a_\mu\exp(\boldsymbol{\phi}') = \exp(\boldsymbol{\phi}')a_\mu$, we only need to consider $a\exp(\phi a^\dagger)|0\rangle$. Following the footnote of [AS10, p. 159], we show that $[a,(a^\dagger)^n] = n\cdot(a^\dagger)^{n-1}$:

$$\begin{aligned}[a,(a^\dagger)n] &= a(a^\dagger)^n - (a^\dagger)^n a = (a(a^\dagger)^{n-1} - (a^\dagger)^{n-1}a)a^\dagger + (a^\dagger)^{n-1} \\ &= [a,(a^\dagger)^{n-1}]a^\dagger + (a^\dagger)^{n-1} = (n-1)(a^\dagger)^{n-1}a^\dagger + (a^\dagger)^{n-1} \\ &= n\cdot(a^\dagger)^{n-1} \ .\end{aligned}$$

With this result, and $a|0\rangle = 0$ we conclude:

$$\begin{aligned}a\exp(\phi a^\dagger)|0\rangle &= a\exp(\phi a^\dagger)|0\rangle - \exp(\phi a^\dagger)a|0\rangle = [a,\exp(\phi a^\dagger)]|0\rangle \\ &= \sum_n \frac{\phi^n}{n!}[a,(a^\dagger)^n]\ |0\rangle = \sum_n \frac{\phi^n}{n!}n\cdot(a^\dagger)^{n-1}|0\rangle \\ &= \phi\sum_{n-1}\frac{\phi^{n-1}(a^\dagger)^{n-1}}{(n-1)!}|0\rangle = \phi\exp(\phi a^\dagger)|0\rangle \ .\end{aligned}$$

ii) For the second property, we observe that:

$$\begin{aligned} a^\dagger \exp(\phi a^\dagger) &= a^\dagger \sum_n \frac{\phi^n (a^\dagger)^n}{n!} = \sum_n \frac{\phi^n (a^\dagger)^{n+1}}{n!} \\ &= \sum_n \frac{1}{n+1} \frac{\partial}{\partial \phi} \frac{\phi^{n+1}(a^\dagger)^{n+1}}{n!} = \frac{\partial}{\partial \phi} \sum_n \frac{\phi^{n+1}(a^\dagger)^{n+1}}{(n+1)!} \\ &= \frac{\partial}{\partial \phi} \sum_n \frac{\phi^{n+1}(a^\dagger)^{n+1}}{(n+1)!} + \frac{\partial}{\partial \phi} 1 = \frac{\partial}{\partial \phi} \sum_n \left(\frac{\phi^{n+1}(a^\dagger)^{n+1}}{(n+1)!} + 1 \right) \\ &= \frac{\partial}{\partial \phi} \sum_{n+1} \frac{\phi^{n+1}(a^\dagger)^{n+1}}{(n+1)!} = \frac{\partial}{\partial \phi} \exp(\phi a^\dagger) \ . \end{aligned}$$

iii) We follow the proof of [AS10, p. 159]. From the first property we infer that $\exp(\sum_\mu z_\mu a_\mu)|\phi\rangle = \exp(\sum_\mu z_\mu \phi_\mu)|\phi\rangle$ for arbitrary $z_\mu \in \mathbb{C}$. Thus:

$$\begin{aligned} \langle \theta | \phi \rangle &= \langle 0 | \exp\left(\sum_\mu \overline{\theta}_\mu a_\mu \right) |\phi\rangle = \langle 0 | \exp\left(\sum_\mu \overline{\theta}_\mu \phi_\mu \right) |\phi\rangle \\ &= \exp\left(\sum_\mu \overline{\theta}_\mu \phi_\mu \right) \langle 0 || \phi \rangle = \exp\left(\sum_\mu \overline{\theta}_\mu \phi_\mu \right) \ . \end{aligned}$$

□

Theorem 6.3.4.
The identity operator on the Fock space can be expressed as follows:

$$\mathbb{1} = \int d(\overline{\phi}, \phi) \ \exp\left(- \sum_\mu \overline{\phi}_\mu \phi_\mu \right) |\phi\rangle\langle\phi| \ ,$$

where $d(\overline{\phi}, \phi) = \prod_j \frac{1}{\pi} d\overline{\phi}_j \ d\phi_j$.

Proof 6.3.5.
This proof follows [AS10, p. 160].

The representation of $\mathcal{W}(\mathcal{H} \oplus \mathcal{H}^*)$ on $S(\mathcal{H})$ is irreducible, such that we want to use a result from Schur's lemma, corollary D.1.10. We need to show that $X = \int_\wedge d(\overline{\phi}, \phi) \ \exp\left(- \sum_\mu \overline{\phi}_\mu \phi_\mu \right) |\phi\rangle\langle\phi|$ commutes with every $a_\mu^\dagger$ and a_μ.

$$\begin{aligned} a_\mu^\dagger X &= \int d(\overline{\phi}, \phi) \ \exp\left(- \sum_\mu \overline{\phi}_\mu \phi_\mu \right) a_\mu^\dagger |\phi\rangle\langle\phi| \\ &= \int d(\overline{\phi}, \phi) \ \left(a_\mu^\dagger |\phi\rangle \right) \langle\phi| \exp\left(- \sum_\mu \overline{\phi}_\mu \phi_\mu \right) \\ &= \int d(\overline{\phi}, \phi) \ \left(\partial_{\phi_\mu} |\phi\rangle \right) \langle\phi| \exp\left(- \sum_\mu \overline{\phi}_\mu \phi_\mu \right) \end{aligned}$$

$$\overset{\text{by parts}}{=} -\int d(\overline{\phi},\phi)\ |\phi\rangle\langle\phi|\partial_{\phi_\mu}\exp\left(-\sum_\mu \overline{\phi}_\mu\phi_\mu\right)$$
$$= -\int d(\overline{\phi},\phi)\ \exp\left(-\sum_\mu \overline{\phi}_\mu\phi_\mu\right) a_\mu^\dagger|\phi\rangle\langle\phi|(-\overline{\phi}_\mu)$$
$$= \int d(\overline{\phi},\phi)\ \exp\left(-\sum_\mu \overline{\phi}_\mu\phi_\mu\right) a_\mu^\dagger|\phi\rangle\langle\phi|a_\mu^\dagger = Xa_\mu^\dagger\ .$$

In the line where we used integration by parts, we used that $\langle\phi| = \langle 0|\exp(\boldsymbol{\phi}^\dagger)$ does not depend on ∂_{ϕ_μ}. The calculation for a_μ is almost the same and can be found in [AS10, p. 160]. From corollary D.1.10 we know that:

$$\int d(\overline{\phi},\phi)\ \exp\left(-\sum_\mu \overline{\phi}_\mu\phi_\mu\right)|\phi\rangle\langle\phi| \sim \mathbb{1}\ .$$

To obtain the proportionality constant we consider $\langle 0|X|0\rangle$, using corollary 6.1.5 for $n = 2$:

$$\langle 0|X|0\rangle = \int d(\overline{\phi},\phi)\ \exp\left(-\sum_\mu \overline{\phi}_\mu\phi_\mu\right)\langle 0|\phi\rangle\langle\phi|0\rangle$$
$$= \prod_j \int \frac{1}{\pi} d\overline{\phi}_j\ d\phi_j \exp\left(-\overline{\phi}_j\phi_j\right) = 1\ .$$

This concludes the proof. □

6.3.2. Functional integral for bosons: The grand canonical partition function

The derivation is the same as for fermions, because of the similarity of the coherent states. In fact, the derivation becomes easier because of commutativity. Some notable difference are however:

$$\mathcal{Z} = \int d(\overline{\phi},\phi)\ \exp\left(-\sum_\mu \overline{\phi}_\mu\phi_\mu\right)\langle\phi|e^{-\beta H}|+\phi\rangle\ ,$$

leading to the condition $\phi_M = \phi_1$. Thus in the continuum limit we obtain again:

$$\boxed{\mathcal{Z} = \int D(\overline{\psi},\psi)\ e^{-\int_0^\beta d\tau\ \overline{\psi}(\tau)\partial_\tau\psi(\tau)+H(\overline{\psi}(\tau),\psi(\tau))}\ ,}$$

yet with the boundary conditions:

$$\psi(\beta) = \psi(0) \qquad \text{and} \qquad \overline{\psi}(\beta) = \overline{\psi}(0)\ .$$

Part III.

Mathematical methods

Used throughout all of these notes, the mathematical concepts introduced in this part could be equally well at the beginning. The separation of the mathematical tools and their application to physical problems is a decision of the author, to provide rigor when needed at the cost of a certain separation, that would not be possible in a lecture course. That said, this part is not, and does not intend to be, as complete or rigorous as detailed mathematical textbooks, as it spans too wide a range of mathematical fields. However it can serve as some glimpse into the proper mathematical theory.

A

Hilbert sapces and Dirac-notation

The foundation of axiomatic quantum mechanics are Hilbert-spaces and operators. In this chapter we cover the basic definitions and properties of Hilbert spaces, operators and transformations. Furthermore the commonly used Dirac notation gets introduced (after the mathematical objects behind the notation have been defined).

A.1. Hilbert spaces

Loosely speaking, Hilbert spaces are a generalization of $\mathbb{R}^n$ for infinite dimensional n. This section mostly follows [KW06, section 10.4].

Definition A.1.1.
Let $(\mathcal{H}, \|\cdot\|)$ be a normed vector space. If every Cauchy sequence is convergent with respect to the norm $\|\cdot\|$, then the vector space is called **Banach space**. If there is a hermitian scalar product $\langle\cdot,\cdot\rangle$ on $\mathcal{H}$ and every Cauchy sequence converges with respect to the $\langle\cdot,\cdot\rangle$-induced norm, the tuple $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ is called **Hilbert space**.

A hermitian scalar product is a map $\langle\cdot,\cdot\rangle\colon \mathcal{H}\times\mathcal{H}\to\mathbb{C}$ that satisfies the following properties:

Sesquilinearity:
$\langle x, y+z\rangle = \langle x,y\rangle + \langle x,z\rangle$, $\langle x+y, z\rangle = \langle x,z\rangle + \langle y,z\rangle$ and $\langle \alpha x, y\rangle = \alpha\cdot\langle x,y\rangle$ for all $x,y,z\in\mathcal{H}$ and $\alpha\in\mathbb{C}$.

Hermitian symmetry:
$\langle x,y\rangle = \overline{\langle y,x\rangle}$ for all $x,y\in\mathcal{H}$.

Positve definition:
$\langle x,x\rangle \geq 0$ and $\langle x,x\rangle = 0 \;\Leftrightarrow\; x=0$ for all $x\in\mathcal{H}$.

The $\langle\cdot,\cdot\rangle$-induced norm is defined as usual by $|x| = \sqrt{\langle x,x\rangle}$. The hermitian scalar product also induces a natural map in the dual space $\mathcal{H}^* = \mathrm{Hom}(\mathcal{H}\to\mathbb{C})$:

$$\ell\colon \mathcal{H}\longrightarrow\mathcal{H}^*,\; x\longmapsto \ell_x\ .$$

This map is bijective, yet not a homomorphism due to the hermitian symmetry, and is defined by

$$\ell_x(y) = \langle x,y\rangle \quad \forall\, x,y\in\mathcal{H}\ .$$

Definition A.1.2 (Dirac notation)**.**
Using **Dirac notation**, elements of $\mathcal{H}$ are denoted by $|x\rangle$ and are called **ket-**

vectors. Elements of the dual space $\vartheta \in \mathcal{H}^*$ are denoted by $\langle\vartheta|$ and are called **bra-vectors**.

This notation convention leads to an illustrative notation for the dual pairing of $\ell_x \in \mathcal{H}^*$ and $y \in \mathcal{H}$:

$$\ell_x(y) \equiv \langle x|(|y\rangle) =: \langle x|y\rangle = \langle x, y\rangle \ .$$

The Dirac notation does not by chance resemble the scalar product, one of the fundamental objects in quantum mechanics. Hence the following notation convention is hardly surprising:

$$|\alpha \cdot x + \beta \cdot y\rangle \coloneqq \alpha \cdot |x\rangle + \beta \cdot |y\rangle \qquad \forall\ x, y \in \mathcal{H},\ \alpha, \beta \in \mathbb{C}\ .$$

The hermitian symmetry has the consequence, that bra-vectors behave slightly differently:

$$\langle \alpha x + \beta y| = \overline{\alpha}\langle x| + \overline{\beta}\langle y| \ .$$

Definition A.1.3.
A **Hilbert basis** in an infinite-dimensional Hilbert space $\mathcal{H}$ is a sequence $(x_n)_{n\in\mathbb{N}_0} \in \mathcal{H}$, that satisfies the following properties:

i) $\langle x_n, x_m\rangle = \delta_{m,n}$.

ii) Let $f \in \mathcal{H}$ and $\langle x_n, f\rangle = 0$ for all $n \in \mathbb{N}_0$, then $f = 0$ holds.

In the finite-dimensional case Hilbert bases are the same as bases. However, in the infinite-dimensional case these concepts are related, but differ considerably. That is, theorems of linear algebra do not carry over, at least not with some effort to show that they do.

Lemma A.1.4 (Continuity of scalar products)**.**
Let $\mathcal{H}$ be a Hilbert space, the following propositions are fulfilled:

i) Let (x_n), (y_n) be convergent sequences, then

$$\left\langle \lim_{n\to\infty} x_n, \lim_{n\to\infty} y_n \right\rangle = \lim_{n\to\infty} \langle x_n, y_n\rangle \ .$$

ii) Let $\sum_{n\in\mathbb{N}_0} x_n$ be a convergent series, then

$$\left\langle \sum_{n\in\mathbb{N}_0} x_n, \quad y \right\rangle = \sum_{n\in\mathbb{N}_0} \langle x_n, y\rangle \quad \forall\ y \in \mathcal{H}\ .$$

Proof A.1.5.
Using the Cauchy-Schwartz inequality shows that

$$|\langle x_n, y_n\rangle - \langle x, y\rangle| = |\langle x_n, y_n - y\rangle + \langle x_n - x, y_n\rangle| \leq |x_n| \cdot |y_n - y| + |y_n| \cdot |x_n - x| \ .$$

The right hand side is a zero sequence and the inequality shows the assertion. The second property is trivial, since a series is just a special case of a sequence.

Lemma A.1.6 (Bessel inequality)**.**
Let (x_n) be an orthonormal sequence in $\mathcal{H}$, i.e. $\langle x_n, x_m\rangle = \delta_{m,n}$, then the ***Bessel inequality*** *is satisfied:*

$$\sum_{n\in\mathbb{N}_0} \langle x_n, f\rangle \leq |f| \ .$$

This lemma allows to prove essential properties of Hilbert bases, that also show how they relate to orthonormal bases:

Theorem A.1.7.
Let $(x_n)_{n\in\mathbb{N}_0}$ be an orthonormal sequence in a Hilbert space $\mathcal{H}$, then the following properties are equivalent:

i) (x_n) *is a Hilbert basis.*

ii) It holds that $\mathcal{H} = \overline{\text{span}\{x_n | n \in \mathbb{N}_0\}}$.

iii) $\forall\ f \in \mathcal{H}: \quad f = \sum_{n\in\mathbb{N}_0} x_n \cdot \langle x_n, f\rangle$.

iv) $\forall\ f, g \in \mathcal{H}: \quad \langle f, g\rangle = \sum_{n\in\mathbb{N}_0} \langle f, x_n\rangle \cdot \langle x_n, g\rangle$.

v) $\forall\ f \in \mathcal{H}: \quad |f|^2 = \sum_{n\in\mathbb{N}_0} \langle x_n, f\rangle^2$.

Due to their equivalence the properties (iii) and (iv) get summarized in the Dirac notation by

$$\mathbb{1} = \sum_{n\in\mathbb{N}_0} |x_n\rangle\langle x_n| \ .$$

The previous theorem states, that this condition is equivalent to (x_n) being a Hilbert basis. The equation is commonly called **Completeness relation** in the physical literature.

Remark A.1.8.
In the next section we will encounter operators. If the operator is bounded, limit and operator may be exchanged, and thus operator and series as well:

$$\begin{aligned} A\sum_{n=0}^{\infty} c_n|n\rangle = A\lim_{N\to\infty}\sum_{n=0}^{N} c_n|n\rangle &= \lim_{N\to\infty} A\sum_{n=0}^{N} c_n|n\rangle \\ = \lim_{N\to\infty}\sum_{n=0}^{N} c_n A|n\rangle &= \sum_{n=0}^{\infty} c_n A|n\rangle \ . \end{aligned}$$

However, this cannot be done, if the operator is unbounded. Yet, for a Hilbert

basis $\{|x_n\rangle\}$, the last theorem (property iv and iii) allows to write:

$$\langle\phi|A|\psi\rangle = \sum_n \langle\phi|A|x_n\rangle\langle x_n|\psi\rangle = \langle\phi|A\sum_n |x_n\rangle\langle x_n|\psi\rangle \ ,$$

for any operator, without the need to exchange operator and series. For that reason $\mathbb{1} = \sum\limits_{n\in\mathbb{N}_0} |x_n\rangle\langle x_n|$ can be misleading.

A.2. Operators

Physically measurable quantities are represented by self adjoint operators in quantum mechanics. Most of the definitions an theorems about operators can be found in [Wer11, section II.1].

Definition A.2.1.
An **operator** T is a continuous map between Hilbert spaces, $T : X \to Y$.

With point-wise addition and scalar multiplication, the space of operators is itself a vector space $\mathrm{Hom}(X \to Y)$. A special subspace is the vector space of continuous operators $\mathcal{L}(X \to Y)$. Continuity is one of the following equivalent conditions

Definition und Lemma A.2.2.
*A linear map T is called **continuous**, if one (and thus all) of the following conditions is fulfilled:*

i) Let (x_n) be a convergent sequence in X with $\lim\limits_{n\to\infty} x_n = x$. Then $\lim\limits_{n\to\infty} Tx_n = Tx$ holds.

ii) $\forall\ x_0 \in X\ \forall \varepsilon > 0 \quad \exists\ \delta > 0\colon\ \|x - x_0\|_X < \delta$
$\Rightarrow\ \|Tx - Tx_0\|_Y < \varepsilon \quad \forall x \in X$.

iii) Let $U \subset Y$ be open in Y, then $T^{-1}(U) \subset X$ is open in X.

Two essential properties of operators are:

Theorem A.2.3.

i) Let $T\colon X \to Y$ be a linear map between normed spaces, then the following claims are equivalent:

- *T is continuous on X.*
- *T is bounded, i.e. $\exists\ K \geq 0\colon\ \|Tx\|_Y \leq K \cdot \|x\|_X \quad \forall\ x \in X$.*
- *T is continuous in $0 \in X$.*

ii) Let $D \subset X$ be a dense subspace of the normed space X, let Y be a Banach space and $T\colon D \to Y$ a bounded operator. Then there is a continuous

extension, i.e. a continuous operator $\widehat{T}\colon X \to Y$ *with* $\widehat{T}\big|_D = T$.

The equivalence of continuity and boundedness allows to define a norm on $\mathcal{L}(X \to Y)$:

Definition A.2.4.
The **operator norm** on $\mathcal{L}(X \to Y)$ is defined by

$$\|T\|_{Op} \coloneqq \inf\{K \geq 0 \mid \|Tx\|_Y \leq K \cdot \|x\|_X \quad \forall\, x \in X\} \ .$$

Alternative ways to define the operator norm are given in the next lemma:

Lemma A.2.5.
It holds that:

$$\|T\|_{Op} = \sup_{x\neq 0} \frac{\|Tx\|_Y}{\|x\|_X} = \sup_{\|x\|_X=1} \|Tx\|_Y = \sup_{\|x\|_X\leq 1} \|Tx\|_Y \ .$$

For the properties of self adjoint operators we follow [Zir13] here. To define self adjoint operators, the adjoint of an operator has to be defined first.

Definition A.2.6.
Let T be a bounded operator. The operator $T^\dagger : Y \to X$ is called the **adjoint operator** of T, if

$$\langle Tx, y\rangle_Y = \langle x, T^\dagger y\rangle_X \quad \forall\, x \in X,\ y \in Y \ .$$

For an unbounded dens operator $T : D(T) \subset X \to Y$ the operator $T^\dagger : D(T^\dagger) \subset Y \to X$, with $D(T^\dagger)$ chosen maximally, is called **formally adjoint operator**, if:

$$\langle Tx, y\rangle_Y = \langle x, T^\dagger y\rangle_X \quad \forall\, x \in D(T),\ y \in D(T^\dagger) \ .$$

Definition A.2.7.
Let $T \in End(\mathcal{H})$, then T is called **self adjoint/hermitian**, if $T^\dagger = T$. The operator T is called **unitary**, if:

$$\langle Tx, Ty\rangle = \langle x, y\rangle \quad \forall x, y \in \mathcal{H} \ .$$

The completeness relation allows for a basis decomposition, with respect to a Hilbert basis, of operators:

$$\begin{aligned} Ty &= \sum_{n\in\mathbb{N}_0} x_n \cdot \langle x_n, Ty\rangle = \sum_{n\in\mathbb{N}_0} x_n \cdot \left\langle x_n, T \sum_{m\in\mathbb{N}_0} x_m \cdot \langle x_m, y\rangle \right\rangle \\ &= \sum_{m,n} x_n \cdot \langle x_n, Tx_m\rangle \cdot \langle x_m, y\rangle = \sum_{m,n} x_n \langle x_n, Tx_m\rangle \cdot \ell_{x_m}(y) \ . \end{aligned}$$

Hence the operator can be expressed by

$$T = \sum_{m,n} x_n \otimes (\langle x_n, Tx_m\rangle \cdot \ell_{x_m}) \ .$$

In the Dirac notation this reads

$$T = \sum_{m,n} |x_n\rangle\langle x_n|Tx_m\rangle\langle x_m| \ .$$

To shorten the notation further, if the Hilbert basis is understood, one simply writes $|n\rangle \coloneqq |x_n\rangle$. Another convention is to write

$$\langle x_n|Tx_m\rangle = \langle x_n|T|x_m\rangle \ .$$

Application of both conventions allows to recognize the completeness relation, that is used in physics, more easily:

$$T = \sum_{m,n} |n\rangle\langle n|T|m\rangle\langle m| = \left(\sum_n |n\rangle\langle n|\right) T \left(\sum_m |m\rangle\langle m|\right) \ „ = “ \ \mathbb{1}T\mathbb{1} \ .$$

The scalar products $\langle n|T|m\rangle$, called **matrix elements**, are the coefficients of the matrix representation w.r.t the Hilbert basis.

Example A.2.8.
The matrix elements of the adjoint operator are:

$$\langle m|A^\dagger|n\rangle = \overline{\langle A^\dagger n|m\rangle} = \overline{\langle n|A|m\rangle} \qquad \Rightarrow \quad A^\dagger = \overline{A^T} = \overline{A}^T \ .$$

For the special case $A = |x\rangle\langle y|$ the adjoint operator has a very simple form:

$$A^\dagger = |y\rangle\langle x| \ .$$

Theorem A.2.9.

i) Eigen values of hermitian operators are real.

ii) The eigen vectors of different eigen values of hermitian operators are orthogonal.

Proof A.2.10.

i)
$$T|x\rangle = \lambda|x\rangle \quad \Rightarrow \quad \langle x|T|x\rangle = \lambda\langle x|x\rangle = \overline{\lambda}\langle x|x\rangle \quad \Rightarrow \quad \overline{\lambda} = \lambda \ .$$

ii)
$$T|x\rangle = \lambda|x\rangle \ , \ T|y\rangle = \varphi|x\rangle \quad \Rightarrow \quad \langle x|T|y\rangle = \lambda\langle x|y\rangle = \varphi\langle x|y\rangle$$
$$\overset{\lambda\neq\varphi}{\Longrightarrow} \quad \langle x|y\rangle = 0 \ .$$

□

A.3. Transformation of operators

In the same way, coordinate transformations can be realized passively, choosing a new chart, or actively by using a diffeomorphism, the dynamics in quantum mechanics can be incorporated in the vectors or in the operators. In the first case, the Hilbert vectors depend on the time as parameters; in the second case, the operators depend on the time. To describe the same physical situation the expectation values, i.e. ultimately the matrix elements $\langle x|T|y\rangle$, have to stay invariant. Consider a transformation of the vectors $U : \mathcal{H} \to \mathcal{H}$, with $|x'\rangle = U|x\rangle$ and $|y'\rangle = U|y\rangle$. The condition for the equivalent transformation of the operator T into T' thus is

$$\langle x'|T|y'\rangle = \langle x|T'|y\rangle \ .$$

Plugging in the definitions of $|x'\rangle$ and $|y'\rangle$ yields

$$\langle x'|T|y'\rangle = \langle Ux|T|Uy\rangle = \langle x|U^\dagger TU|y\rangle \qquad \Rightarrow \quad T' = U^\dagger TU \ .$$

Example A.3.1 (Change of basis)**.**
Let $(|\psi_n\rangle)$ and $(|\varphi_n\rangle)$ be two Hilbert bases of $\mathcal{H}$. In the finite-dimensional case a change of basis would mean to find coefficients $c_{m,n}$, such that

$$|\psi_n\rangle = \sum_m c_{m,n}|\varphi_m\rangle \ .$$

The completeness relation allows to realize this equation by defining $c_{m,n} = \langle\varphi_m|\psi_n\rangle$:

$$|\psi_n\rangle = \mathbb{1}|\psi_n\rangle = \sum_m |\varphi_m\rangle\langle\varphi_m|\psi_n\rangle = \sum_m \langle\varphi_m|\psi_n\rangle \cdot |\varphi_m\rangle \ .$$

In the finite-dimensional case, the coefficients $c_{m,n}$ are interpreted as a basis transformation matrix C. This matrix is the matrix representation of the basis change operator $\widehat{C}$ in the basis $(|\varphi_m\rangle)$. Let k be the dimension in the finite-dimensional case, then it folows that

$$\widehat{C}|\varphi_n\rangle = C\begin{pmatrix}1\\0\\\vdots\end{pmatrix} = \begin{pmatrix}c_{1,n}\\\vdots\\c_{k,n}\end{pmatrix} = \sum_{m=1}^{k} c_{m,n}|\varphi_m\rangle = |\psi_n\rangle \ .$$

In the infinite-dimensional case the basis transformation operator can be defined by

$$1)\ \widehat{C}|\varphi_n\rangle = |\psi_n\rangle \qquad 2)\ \langle\varphi_m|\widehat{C}|\varphi_n\rangle = c_{m,n} = \langle\varphi_m|\psi_n\rangle \ .$$

In the finite-dimensional as well as in the infinite-dimensional case, the operator $\widehat{C}$ is given by

$$\widehat{C} = \sum_k |\psi_k\rangle\langle\varphi_k| \ .$$

This means that the basis transformation operator is unitary:

$$\widehat{C}^{-1} = \sum_k |\varphi_k\rangle\langle\psi_k| = \widehat{C}^\dagger \qquad \Leftrightarrow \qquad \langle\widehat{C}x, \widehat{C}y\rangle = \langle x, \widehat{C}^\dagger\widehat{C}y\rangle = \langle x, y\rangle \ .$$

Now, let T be an operator that shall be transformed with $\widehat{C}$. As before we use

$$T' = \widehat{C}^\dagger T \widehat{C} \ .$$

Since the basis transformation operator us unitary, we recover the transformation law known from linear algebra:

$$T' = \widehat{C}^\dagger T \widehat{C} = \widehat{C}^{-1} T \widehat{C} \ .$$

B

Generators, rotation groups and vector operators

Rotations belong to the central symmetry transformation of physics. By Noether's theorem the angular momentum is tied to this symmetry transformation allowing to define angular momentum in quantum mechanics. Furthermore, the rotation group allows to define vector like quantities by analogies to $\mathbb{R}^3$. The mathematically foundation of rotation groups is the theory of Lie groups.

B.1. Lie groups and Lie algebras

In this section, we cover the basics of Lie group theory necessary to understand angular momentum. This section is greatly inspired by [Sch09] and follows it closely in some parts.

B.1.1. Definitions

Definition B.1.1.
A **Lie group** is a differentiable manifold G together with a group structure $\circ\colon G \times G \to G$, such that the operation $\circ$ and the inversion $g \mapsto g^{-1}$ are differentiable.

Let V be a vector space. A map $[\cdot,\cdot] : V \times V \to V$ is called **Lie bracket** if it satisfies the following properties:

Bilinearity

$$[\alpha x + \beta y, z] = \alpha[x,z] + \beta[y,z]$$
$$\text{and} \quad [x, \alpha y + \beta z] = \alpha[x,y] + \beta[x,z] \quad \forall\ x,y,z \in V\ \alpha,\beta \in \mathbb{K}$$

Jacobi-identity

$$[x,[y,z]] + [y,[z,x]] + [z,[x,y]] = 0 \quad \forall\ x,y,z \in V$$

Alternation

$$[x,y] = -[y,x]$$

Definition B.1.2.
A vector space together with Lie bracket is called **Lie algebra**.

A common example is the vector space of endomorphisms $\mathrm{End}(\mathcal{H})$ together with the commutator $[A,B] = A \circ B - B \circ A$.

Theorem B.1.3.
The general Baker-Campbell-Hausdorff formula, for arbitrary A, B reads

$$e^A e^B = e^{A+B+\frac{1}{2}[A,B]+\dots} ,$$

where we have only written the first two summands. The following summands are combinations of $[A,B]_n$ and $[B,A]_n$.

Lemma B.1.4 (Useful relations for commutators, [Sch13, p. 26])**.**
The commutator on $End(\mathcal{H})$ has the following properties

i) $[A,B] = 0 \quad \Rightarrow \quad e^A \cdot e^B = e^{A+B}$.

ii) Let $[A,B]_{m+1} = [A,[A,B]_m]$ *with* $[A,B]_0 = B$ *then*
$e^A B e^{-A} = \sum_{n=0}^{\infty} \frac{1}{n!}[A,B]_n$

iii) If $[A,B]_2 = [B,A]_2 = 0$, *then the following holds:* $e^{A+B} = e^A \cdot e^B \cdot e^{-\frac{1}{2}[A,B]}$.

Remark B.1.5.
Property iii) is just a consequence of theorem B.1.3. If $[A,B]_2 = [B,A]_2 = 0$ then all other higher commutators vanish and especially $[A+B,[A,B]] = 0$, allowing to write $e^{A+B+\frac{1}{2}[A,B]} = e^{A+B} e^{\frac{1}{2}[A,B]}$.

A very important property of the commutator is the existence of joint eigenvalues:

Theorem B.1.6 (Compatibility Theorem)**.**
Let $A, B \in End(\mathcal{H})$ be two operators. If $[A,B] = 0$, then there exists a joint eigen basis $(|\psi_n\rangle)_n$ of A and B. Conversely, if there is a joint an eigen Hilbert basis of A and B, then $[A,B] = 0$

Proof B.1.7.
The proof can be found on [Wik18]. Here we will not show the general case of degenerate eigen values, but focus on the principle of the proof.

Let $|\psi_n\rangle$ be the single eigen vector of A for the eigenvalue a_n. Since by assumption the operators commute, it holds that

$$A(B|\psi_n\rangle) = B(A|\psi_n\rangle) = a_n \cdot B|\psi_n\rangle .$$

this means, that $B|\psi_n\rangle$ is also an eigen vector of A to the eigen value a_n. In the non-degenerate case, the eigen space in one-dimensional. Hence there is a constant b_n, such that $B|\psi_n\rangle = b_n \cdot |\psi_n\rangle$. Thus $(|\psi_n\rangle)_n$ defines an eigen Hilbert basis of B with eigen values $\{b_n\}$. In the degenerate case, the joint eigen basis needs further construction, such that not all eigen bases of A are also eigen bases of B.

Conversely, assume $(|\psi_n\rangle)_n$ to be a joint eigen Hilbert basis of A and B, then

$$AB|\psi_n\rangle = b_n \cdot A|\psi_n\rangle = b_n \cdot a_n \cdot |\psi_n\rangle = BA|\psi_n\rangle$$
$$\Rightarrow \qquad (AB - BA)|\psi_n\rangle = 0 \ .$$

□

Definition B.1.8.
Let $\mathfrak{g}_1$ and $\mathfrak{g}_2$ be two Lie algebras with Lie brackets $[\cdot,\cdot]_1$ and $[\cdot,\cdot]_2$. A **Lie algebra homomorphism** is a linear map (with respect to the field of the vector space) $\varphi\colon \mathfrak{g}_1 \to \mathfrak{g}_2$, such that:

$$\varphi\left([X,Y]_1\right) = [\varphi(X),\varphi(Y)]_2 \qquad \forall\ X,Y \in \mathfrak{g}_1 \ .$$

Definition B.1.9.
Let G be a Lie group and $g \in G$, then the following diffeomorphisms

$$L_g\colon G \to G,\ x \mapsto g \circ x \equiv gx \quad \text{und} \quad R_g\colon G \to G,\ x \mapsto x \circ g \equiv xg$$

are called **left translation** and **right translationright translation** respectively.

B.1.2. Lie algebra of a Lie group

Let $\mathfrak{X}(G)$ denote the set of smooth vector fields on G. On smooth manifolds, tangent vectors define derivations, such that the vector field commutator can be defined as usual:

$$[v,w](f) = (vw)(f) - (wv)(f) \coloneqq v(w(f)) - w(v(f)) \ .$$

The following calculation shows, that the vector field commutator is the same as the Lie derivative:

$$\begin{aligned}
\left(\mathcal{L}_u v\right)(f) &= \lim_{t\to 0} \frac{(\Phi_{-t})_* v - (\Phi_0)_* v}{t}(f) = \lim_{t\to 0} \frac{(\Phi_{-t})_* v - v}{t}(f) \\
&= \lim_{t\to 0} \frac{1}{t}\left((\Phi_{-t})_* v(f) - v(f)\right) \\
&= \lim_{t\to 0} \frac{1}{t}(v(\varphi_{-t}\circ f) - v(f)) \\
&= \lim_{t\to 0} \frac{1}{t}(v(\varphi_{-t}\circ f) - v(f) + \{\Phi_t^* v(\varphi_{-t}\circ f)\} - \{\Phi_t^* v(\varphi_{-t}\circ f)\}) \\
&= \lim_{t\to 0} \Phi_t^* \left\{ v\left(\frac{\Phi_{-t}\circ f - f}{t}\right)\right\} + \lim_{t\to 0} \frac{\Phi_{-t}^*\{v(f)\} - v(f)}{t} \\
&= \lim_{t\to 0} v\left(\frac{\Phi_{-t}\circ f - f}{t}\right) + \lim_{t\to 0} \frac{\Phi_{-t}^*\{v(f)\} - v(f)}{t} \\
&= v(\mathcal{L}_{-u} f) + \mathcal{L}_u(v(f)) = u(v(f)) - v(u(f)) = [u,v](f) \ .
\end{aligned}$$

Lemma B.1.10.
Let $\phi\colon G_1 \to G_2$ be a differentiable map and let X, Y be smooth vector fields in $\mathfrak{X}(G_1)$, then

$$\phi_*([X,Y]) = [\phi_* X, \phi_* Y] \ .$$

Proof B.1.11.
By linearity it is enough to show $\phi_*(X \circ Y) = \phi_*(X) \circ \phi_*(Y)$:

$$\begin{aligned}\phi_*(X \circ Y)(f) &= (X \circ Y)(f \circ \phi) = X(Y(f \circ \phi)) = X(\phi_* Y(f) \circ \phi) \\ &= (\phi_* X)((\phi_* Y)(f)) = (\phi_* X \circ \phi_* Y)(f) \ .\end{aligned}$$

□

Remark B.1.12.
In the last proof we used the extension of push forwards for vector fields. In general, the push forward (or differential) is a map

$$\phi_*\colon T_p G_1 \longrightarrow T_{\phi(p)} G_2, \quad v \longmapsto \phi_* v \ .$$

In the case of vector fields, i.e. with variable base point, this means:

$$X_p \longmapsto (\phi_* X)_{\phi(p)} = (\phi_* X \circ \phi)_p \ .$$

A different notation for the push forward, we will also use, of a vector field in the point p is $D_p\phi$.

It has been shown so far, that $(\mathfrak{X}(G), [\bullet, \bullet])$ defines a Lie algebra. However this is not what is meant by Lie algebra of a Lie group. Instead one uses a sub-algebra:

Definition B.1.13.
Let L_g be a left translation. A vector field $X \in \mathfrak{X}(G)$ is called **left invariant**, if

$$(D_p L_g) X_p = X_{L_g(p)} \ \forall \ p \in G \ .$$

The terminology has the following reason: Using the definition of the push forward, understanding tangent vectors as derivations, one can write (only for left invariant fields)

$$((L_g)_* X)(f) = X(f \circ L_g) = X(f) \circ L_g \ .$$

A left translation of a left invariant vector field results in the tangent vector of the vector field at the image point of the left translation. Thus the vector field has not changed.

Theorem and definition B.1.14.
The set of left invariant vector fields, together with the vector field commutator,

is called ***the Lie algebra $\mathfrak{g}$ of the Lie group G***

Proof B.1.15.
It is enough to show, that if X and Y are left invariant, so is $[X, Y]$.

$$XY(f \circ L_g) = X(Y(f \circ L_g)) = X(Y(f) \circ L_g) = X(Y(f)) \circ L_g = XY(f) \circ L_g \ ,$$

$$\text{and similarly:} \qquad YX(f \circ L_g) = YX(f) \circ L_g \ .$$

Subtraction shows:

$$(XY - YX)(f \circ L_g) = (XY - YX)(f) \circ L_g \qquad \Rightarrow \qquad [X, Y] \in \mathfrak{g} \ .$$

□

Lemma B.1.16.
The Lie algebra $\mathfrak{g}$ of a Lie group G is isomorphic to the tangent space in the one-element:

$$\mathfrak{g} \cong T_eG \ .$$

Proof B.1.17.
Let $X \in \mathfrak{g}$ be left invariant and $g \in G$ arbitrary chosen, then:

$$X_g(f) = X_{L_g(e)}(f) = X_e(f \circ L_g) = \underbrace{(D_eL_gX_e)}_{\in T_gG}(f) \qquad \Rightarrow \qquad X_e \in T_eG \ .$$

The isomorphisms are:

$$\mathfrak{g} \xrightarrow{\cong} T_eG, \ X \longmapsto X_e \qquad \text{and} \qquad T_eG \xrightarrow{\cong} \mathfrak{g}, \ X_e \longmapsto D_eL_gX_e \ .$$

The commutator is carried over by $[X, Y] = [X_e, Y_e]$. □

B.1.3. Exponential map

Definition B.1.18.
Let $(G, \cdot)$ be a Lie group. A **continuous one parameter group** $\{\Phi_t\}_t \subset G$ is a group together with a continuous group homomorphism:

$$(\mathbb{R}, +) \xrightarrow{Hom} (G, \cdot) \ .$$

Theorem B.1.19.
Let $X \in \mathfrak{g}$ be a left invariant vector field and Φ_X the associate integral curve through $e \in G$, i.e. $\frac{d}{dt}\Phi_X(t) = X_{\Phi_X(t)}$ with $\Phi_X(0) = e$, then the following claims

hold:

i) $\Phi_X(t)$ *is defined on* $\mathbb{R}$.

ii) $\Phi_X(s+t) = \Phi_X(s) \cdot \Phi_X(t) \quad \forall\ s,t \in \mathbb{R}$.

iii) $\Phi_{(s\cdot X)}(t) = \Phi_X(s \cdot t) \quad \forall\ s,t \in \mathbb{R}$.

Proof B.1.20 (from [Bau09, proof of theorem 1.2])**.**

ii) Let Φ_X be defined on $I = (t_m, t_M) \subset \mathbb{R}$. Choose a fixed $s \in I$ and define the curved

$$f\colon I \ni \tau \longmapsto \Phi_X(s) \cdot \Phi_X(\tau) \in G$$
$$h\colon (t_m - s, t_M - s) \ni \tau \longmapsto \Phi_X(\tau + s) \in G\ .$$

Both of these maps define smooth curves, that are integral curves of X through $\Phi_X(s)$, i.e. $f(0) = \Phi_X(s) = g(0)$ and

$$\frac{d}{d\tau} f(\tau) = D_{\Phi_X(\tau)} L_{\Phi_X(s)} \left(\frac{d}{d\tau} \Phi_X(\tau) \right) = D_{\Phi_X(\tau)} L_{\Phi_X(s)} (X_{\Phi_X(\tau)})$$
$$= X_{\Phi_X(\tau)\cdot\Phi_X(s)} = X_{f(\tau)} \frac{d}{d\tau} h(\tau) = X_{\Phi_X(s+\tau)} = X_{h(\tau)}\ .$$

Since integral curves are unique, f and h have to agree on their common domain $I \cap (t_m - s, t_M - s)$, such that

$$\Phi_X(s) * \Phi_X(t) = f(t) = h(t) = \Phi_X(s+t)$$

for all $t \in I \cap (t_m - s, t_M - s)$. The extension to all of $\mathbb{R}$ follows from the next property.

i) Assume (t_m, t_M) to be the maximal interval, that Φ_x is defined on, with $t_M < \infty$ and let $\alpha = \min(t_M, |t_m|)$. Define the map f by $f(s) = \Phi_X(\frac{\alpha}{2}) \cdot \Phi_X(s - \frac{\alpha}{2})$. Then,

$$f(0) = \Phi_X(\tfrac{\alpha}{2}) \cdot \Phi_X(-\tfrac{\alpha}{2}) = \Phi_X(0) = e$$

and with $g \coloneqq \Phi_X(\frac{\alpha}{2})$ also

$$\frac{d}{ds} f(s) = D_{\Phi_X(\frac{\alpha}{2})} L_g \left(\frac{d}{ds} \Phi_X(s - \tfrac{\alpha}{2}) \right) = X_{\Phi_X(\frac{\alpha}{2})\cdot\Phi_X(s-\frac{\alpha}{2})} = X_{f(s)}\ .$$

But this means, $f(s)$ is an integral curve of X through e, that extends Φ_X beyond the assumed maximal interval I. Similarly it an extension beyond the lower bond t_m can be shown, such that $I = \mathbb{R}$.

iii) Consider the map $f(t) = \Phi_X(s \cdot t)$. Then $f(0) = e$ and $\frac{d}{dt} f(t) = sX_{f(s)}$. Thus $f(t)$ is an integral curve of sX through e, which means $\Phi_X(s \cdot t) = \Phi_{sX}(t)$.

□

This theorem is the motivation to define the exponential map (of Lie group theory):

Definition B.1.21.
The **exponential map** is defined by:

$$\exp\colon \mathfrak{g} \longrightarrow G,\ X \longmapsto \exp(X) \equiv \Phi_X(1)\ .$$

Lemma B.1.22.
Let 0 *be the zero vector field of* $\mathfrak{g}$*, then:*

$$D_0 \exp = \mathbb{1}_{\mathfrak{g}}\ .$$

Proof B.1.23.
First we will show, that $D_0 \exp$ is a map in $\mathrm{End}(\mathfrak{g})$:

$$D_0 \exp : T_0\mathfrak{g} \cong \mathfrak{g} \longrightarrow T_{\exp(0)}G = T_eG \cong \mathfrak{g}\ .$$

The second isomorphism has been shown in lemma B.1.16. The first isomorphism is the general isomorphism between vector spaces, treated as manifolds, and their tangent spaces. Since differential are linear by definition, the claim $D_0 \exp \in \mathrm{End}(\mathfrak{g})$ follows.

Let $X \in \mathfrak{g}$, then $\gamma(t) = t \cdot X$ describes a curve in $\mathfrak{g}$ with the properties $\gamma(0) = 0$ and $\gamma'(0) = X$. For the differential it follows, that:

$$\begin{aligned}(D_0 \exp)(X) &= \left.\frac{d}{dt}\right|_{t=0} (\exp(0 + t \cdot X)) = \left.\frac{d}{dt}\right|_{t=0} \Phi_{t\cdot X}(1) \\ &= \left.\frac{d}{dt}\right|_{t=0} \Phi_X(t) = X_e \equiv X\ .\end{aligned}$$

□

A direct consequence of the previous lemma is

$$\left.\frac{d}{dt}\right|_{t=0} \exp(tX) = (D_0 \exp)X = X\ .$$

Corollary B.1.24 (Properties of the exponential map)**.**
Let $s, t \in \mathbb{R}$ *and* $X \in \mathfrak{g}$*, then:*

i) $\exp(0) = e$.
ii) $\exp(-X) = (\exp(X))^{-1}$.
iii) $\exp((t+s)X) = \exp(tX) \circ \exp(sX)$.

Theorem and definition B.1.25.
Every one parameter sub group $\{g_t\}_{t\in\mathbb{R}}$ *of* G *is generated by an element of* $\mathfrak{g}$*, i.e.*

there is an $X \in \mathfrak{g}$, such that

$$\{g_t\}_{t\in\mathbb{R}} = \{\exp(tX)\}_{t\in\mathbb{R}} \ .$$

Such an element is called ***generator****.*

Remark B.1.26.
If the Lie algebra is a matrix algebra, that is closed under multiplication, then the exponential map is the well known e-function:

$$\exp(A) = e^A := \sum_{n=0}^{\infty} \frac{A^n}{n!} \ .$$

This can be seem as follows: If e^A exists, so does $f(t) = e^{tA}$. With $f(0) = \mathbb{1}$ and

$$\frac{d}{dt} f(t) = e^{tA} A \equiv e^{tA} A_{tA} = (D_{tA} L_{e^{tA}}) A_{tA} = A_{e^{tA}} = A_{f(t)} \ ,$$

we see that $f(t)$ is the integral curve of A through $\mathbb{1}$, i.e. $\Phi_A(t) = f(t)$. Thus:

$$\exp(A) = \Phi_A(1) = e^{1\cdot A} = e^A \ .$$

B.2. Rotation group and vector operators

A part of physical quantities, known from classical physics, are vectors in $\mathbb{R}^3$. However, the meaning of a vector in this context is not the mathematical one, i.e. being element of a vector space, but the geometrical one. This discrepancy leads to the necessity to define vector operators, to recover the geometrical properties of vector like quantities. In this section we will briefly cover the concepts of rotation groups and vector operators, as can be found in [Zir13] and [RW08].

B.2.1. Rotation group and its generators

A one parameter group of unitary operators $\{U_t\}_{t\in\mathbb{R}}$ is called **strongly continuous group**, if for all $t_0 \in \mathbb{R}$ the following limit

$$\lim_{t\to t_0} U_t = U_{t_0}$$

is valid, as well as the condition that U_t defines a local flow:

$$U_{s+t} = U_s \circ U_t \ .$$

The concept of strongly continuous groups allows to formulate a very important theorem for the connection between generators and symmetries:

Theorem B.2.1 (Stone's theorem)**.**
Let $\{U_t\}_{t\in\mathbb{R}}$ be a one parameter group of unitary operators that is strongly continuous, then the following is true:

i) For every U_t there exists a hermitian generator, i.e. a hermitian operator A such that $U_t = e^{itA}$.

ii) Let A be a hermitian operator. Then this operator generates a one parameter group of unitary operators : $\{e^{itA}\}_{t\in\mathbb{R}}$.

In consideration of the Schrödinger-equation $i\hbar\partial_t|\psi(t)\rangle = H|\psi(t)\rangle$, generators are defined with an additional coefficient $-i/\hbar$:

$$U_t = e^{-\frac{i}{\hbar}tB} \ .$$

For the rest of this subsection we will denote the standard scalar product of $\mathbb{R}^n$ with $\langle\cdot,\cdot\rangle$.

Definition B.2.2.
The **rotation group** $\mathrm{SO}(n)$ is defined by

$$\mathrm{SO}(n) \coloneqq \{R : \mathbb{R}^n \to \mathbb{R}^n | \langle Rx, Ry\rangle = \langle x, y\rangle \ \forall\ x, y \in \mathbb{R}^n, \det(R) = 1\} \ .$$

The maps $\mathbf{R}$ are called **rotations**. With $R_{\vec{\alpha}}$ we denote a rotation around the axis $\vec{e}_\alpha$ with rotation angle $|\vec{\alpha}|$.

Definition B.2.3.
The space of skew-symmetric maps

$$\mathfrak{so}(V) \coloneqq \{A : V \to V | A^T = -A\} \ ,$$

together with the commutator is a Lie algebra. For the special case $V = \mathbb{R}^n$ it is denoted by $\mathfrak{so}(n)$.

It can be shown, that $\mathrm{SO}(n)$ is a Lie group of dimension $\frac{n(n-1)}{2}$ with Lie algebra $\mathfrak{so}(n)$. In the special case of $n = 3$, the dimensions of rotations and space agree, allowing an Isomorphism between $\mathfrak{so}(n)$ and $(\mathbb{R}^3, \times)$, often used in introductory books about classical mechanics.

Lemma B.2.4.
A rotation R_φ with angle φ around the x_j-axis has the following generator in $\mathfrak{so}(3)$:

$$J_i = -\sum_{j,k} \varepsilon_{ijk} \vec{e}_j \otimes \vec{e}^{\,*}_k \ .$$

Here $\{\vec{e}_i\}$ denotes the standard basis of $\mathbb{R}^3$ with dual basis $\{\vec{e}^{\,*}_i\}$.

Proof B.2.5.

Œ for $j = 3$ we consider a rotation around the z-axis:

$$R_\varphi = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) & 0 \\ \sin(\varphi) & \cos(\varphi) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

For $R_\varphi = e^{\varphi J_3}$ differentiation at $\varphi = 0$ yields:

$$J_3 = \frac{d}{d\varphi}\bigg|_{\varphi=0} e^{\varphi J_3} = \frac{d}{d\varphi}\bigg|_{\varphi=0} R_\varphi \ .$$

Hence:

$$J_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \vec{e}_2 \otimes \vec{e}^{\,*}_1 - \vec{e}_1 \otimes \vec{e}^{\,*}_2 \ .$$

The remaining axes can be shown similarly. □

The proof allows to read of the Lie algebra of SO(3) easily, by checking the dimensions:

Corollary B.2.6.
The Lie algebra of the Lie group SO(3) *is* $\mathfrak{so}(3)$.

A straightforward calculation results in the following lemma:

Lemma B.2.7.
The generators of the $\mathfrak{so}(3)$ *satisfy the following commutator relation:*

$$[J_i, J_j] = \varepsilon_{ijk} J_k \qquad \Big[\sum_{j=1}^{3} J_j^2, J_i\Big] = 0 \ .$$

With the quantum mechanical convention for generators, we find

$$L_x = i\hbar J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i\hbar \\ 0 & i\hbar & 0 \end{pmatrix} \qquad L_y = i\hbar J_2 = \begin{pmatrix} 0 & 0 & i\hbar \\ 0 & 0 & 0 \\ -i\hbar & 0 & 0 \end{pmatrix}$$

$$L_z = i\hbar J_3 = \begin{pmatrix} 0 & -i\hbar & 0 \\ i\hbar & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \ .$$

The commutator relations assume the following form:

$$[L_\alpha, L_\beta] = (i\hbar)^2 \varepsilon_{\alpha\beta\gamma} L_\gamma \qquad \left[\sum_{\alpha=1}^{3} L_\alpha^2, L_\beta\right] = 0 \ .$$

Theorem B.2.8.

A rotation around $\vec{\alpha}$ with rotation angle $|\vec{\alpha}|$ has the following generator

$$J = \frac{1}{|\vec{\alpha}|} \sum_{i=1}^{3} \alpha_i \cdot J_i \ \in \mathfrak{so}(3) \ .$$

Thus the rotation map has the following form: $R_{|\vec{\alpha}|} = e^{|\vec{\alpha}|J} := e^{J_{|\vec{\alpha}|}}$.

Proof B.2.9.
Theorem B.1.25 allows us to deduce, that $R_{|\vec{\alpha}|}$ is an element of SO(3). It remains to show that is has the eigen vector $\vec{\alpha}$ to the eigen value 1, and that the rotation angle is $|\vec{\alpha}|$.

A calculation shows

$$J_{|\vec{\alpha}|}^2 = -|\vec{\alpha}|^2 J_{|\vec{\alpha}|} \ .$$

This equation can be extended to

$$J_{|\vec{\alpha}|}^{2n+1} = (-1)^n |\vec{\alpha}|^{2n} J_{|\vec{\alpha}|} \quad \text{and} \quad J_{|\vec{\alpha}|}^{2n+2} = (-1)^n |\vec{\alpha}|^{2n} J_{|\vec{\alpha}|}^2 \ .$$

Separating the exponential series in sine and cosine yields

$$\begin{aligned}
R_{|\vec{\alpha}|} = e^{J_{|\vec{\alpha}|}} &= \sum_{n\in\mathbb{N}_0} \frac{1}{n!} J_{|\vec{\alpha}|}^n = J_{|\vec{\alpha}|}^0 + \sum_{n\in 2\mathbb{N}_0+1} \frac{1}{n!} J_{|\vec{\alpha}|}^n + \sum_{n\in 2\mathbb{N}} \frac{1}{n!} J_{|\vec{\alpha}|}^n \\
&= \mathbb{1} + \sum_{n\in\mathbb{N}_0} \frac{1}{(2n+1)!} J_{|\vec{\alpha}|}^{2n+1} + \sum_{n\in\mathbb{N}_0} \frac{1}{(2n+2)!} J_{|\vec{\alpha}|}^{2n+2} \\
&= \mathbb{1} + \sum_{n\in\mathbb{N}_0} \frac{(-1)^n \|\vec{\alpha}\|^{2n}}{(2n+1)!} J_{|\vec{\alpha}|} + \sum_{n\in\mathbb{N}_0} \frac{(-1)^n \|\vec{\alpha}\|^{2n}}{(2n+2)!} J_{|\vec{\alpha}|}^2 \\
&= \mathbb{1} + \frac{\sin(|\vec{\alpha}|)}{|\vec{\alpha}|} J_{|\vec{\alpha}|} + \frac{1-\cos(|\vec{\alpha}|)}{|\vec{\alpha}|^2} J_{|\vec{\alpha}|}^2 \\
&= \mathbb{1} + \sin(|\vec{\alpha}|) J + (1-\cos(|\vec{\alpha}|)) J^2 \ .
\end{aligned}$$

One can calculate that $J\vec{\alpha} = 0$. Hence $R_{|\vec{\alpha}|}\vec{\alpha} = \vec{\alpha}$. The third property of corollary B.1.24 assures that the rotation angle scales linearly with $|\vec{\alpha}|$. So it is enough to check the rotation angle for only one value. A short calculation shows that $R_{|\vec{\alpha}|}$ for $|\vec{\alpha}| = \pi/2$ is indeed the rotation with rotation angle $\pi/2$. □

B.2.2. Vector operators

As announced in the beginning of this section, we will use the rotation group SO(3) to define vector operators here.

Definition B.2.10.
Let $\mathbf{A} = (A_1, A_2, A_3)$ be a set of operators $A_1 : \mathcal{H} \to \mathcal{H}$. Then $\mathbf{A}$ is called **vector**

operator, if the operators transform like a vector in $\mathbb{R}^3$ for rotations. That is:

$$R\begin{pmatrix}A_1\\A_2\\A_3\end{pmatrix} = \begin{pmatrix}A_1'\\A_2'\\A_3'\end{pmatrix} \quad \text{and} \quad A_i' = \widehat{R}^\dagger A \widehat{R} \ .$$

Here R is a rotation and $\widehat{R}$ its representation on the Hilbert space.

This definition uses the geometrical definition of vectors in $\mathbb{R}^3$. However the condition to be a vector operator can be bothersome to check like this. The next lemma gives an alternative definition:

Lemma B.2.11.
A *is a vector operator, if and only if the commutator relations* $[L_\alpha, A_\beta] = i\hbar\varepsilon_{\alpha\beta\gamma}A_\gamma$ *are satisfied.*[1]

Proof B.2.12 (Œ for rotations around the $\boldsymbol{z}$-axis and $\boldsymbol{A_1}$)**.**
The first thing to notice, is the equality of $\frac{d}{d\varphi}\Big|_{\varphi=0} A_1'(\varphi)$ and $\frac{i}{\hbar}[L_3, A_1]$:

$$\begin{aligned}\tfrac{i}{\hbar}[L_3, A_1] &= \frac{i}{\hbar}(L_3A_1 - A_1L_3)\\ &= \left(\frac{d}{d\varphi}e^{\frac{i}{\hbar}\varphi L_3}A_1\right)e^{-\frac{i}{\hbar}\varphi L_3}\Big|_{\varphi=0} + e^{\frac{i}{\hbar}\varphi L_3}\left(\frac{d}{d\varphi}A_1e^{-\frac{i}{\hbar}\varphi L_3}\right)\Big|_{\varphi=0}\\ &= \frac{d}{d\varphi}\Big|_{\varphi=0} e^{\frac{i}{\hbar}\varphi L_3}A_1e^{-\frac{i}{\hbar}\varphi L_3} = \frac{d}{d\varphi}\Big|_{\varphi=0} A_1'(\varphi) \ .\end{aligned}$$

Using the rotation matrix we find:

$$(R_\varphi\mathbf{A})_1 = \cos(\varphi)A_1 - \sin(\varphi)A_2 \qquad (R_\varphi\mathbf{A})_2 = \sin(\varphi)A_1 + \cos(\varphi)A_2$$
$$(R_\varphi\mathbf{A})_3 = A_3 \ .$$

Vector operator ⇒ commutator relation:
If **A** is a vector operator, then $(R_\varphi\mathbf{A})_1 = A_1'(\varphi)$. Differentiation at $\varphi = 0$ shows:

$$-A_2 = \frac{d}{d\varphi}\Big|_{\varphi=0}(R_\varphi\mathbf{A})_1 = \frac{d}{d\varphi}\Big|_{\varphi=0} A_1'(\varphi) = \tfrac{i}{\hbar}[L_3, A_1]$$
$$\Leftrightarrow \quad [L_3, A_1] = i\hbar A_2 \ .$$

Commutator relation ⇒ vector operator:
We have to show, that $(R_\varphi\mathbf{A})_1 = A_1'(\varphi)$ holds. From the commutator relations it follows that

$$\frac{d}{d\varphi}\Big|_{\varphi=0} A_1'(\varphi) = \tfrac{i}{\hbar}[L_3, A_1] = -A_2 = \frac{d}{d\varphi}\Big|_{\varphi=0}(R_\varphi\mathbf{A})_1 \ .$$

[1] The reason we use the quantum mechanical definition of generators here is, that the angular momentum operators $\widehat{L}_\alpha$ will be the representations of these generators in properly chosen coordinates.

The last equality can be calculated directly. Hence we find equality up to a constant (with respect to φ):

$$(R_\varphi \mathbf{A})_1 = A_1'(\varphi) + B \ .$$

Chsooing $\varphi = 0$, we find $B \equiv 0$, which proves the claim.

□

Corollary B.2.13.
The generators $\mathbf{L} = (L_1, L_2, L_3)$ *define a vector operator.*

B.2.3. Pauli-matrices

At this point a short digression on $SU(2)$ and the Pauli matrices seems appropriate.

Definition B.2.14.
The Lie group $SU(2)$ is defined by

$$SU(2) \coloneqq \{R : \mathbb{C}^2 \to \mathbb{C}^2 | \langle Rx, Ry\rangle = \langle x, y\rangle \ \forall \ x, y \in \mathbb{C}^2, \quad \det(R) = 1\} \ .$$

The Lie algebra of $SU(2)$ is $\mathfrak{su}(2)$ defined by

$$\mathfrak{su}(2) \coloneqq \{A : \mathbb{C}^2 \to \mathbb{C}^2 | A^T = -A\} \ .$$

Inspecting these matrices reveals that $\mathfrak{su}(2)$ is a 3-dim vector space with the following special basis:

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

These matrices are called **Pauli matrices**. The choice of indices and especially of this basis is absolutely arbitrary. However it is the predominant convention in the physical literature. These matrices have the following commutator relations:

$$[\sigma_j, \sigma_k] = 2i\varepsilon_{jkl}\sigma_l \ .$$

B.3. Eigen states of angular momentum operators

A vector operator of particular interest is the angular momentum operator. For that reason, the eigen states of angular momentum operators (or in general the representations of generators of rotations) will be inquired.

B.3.1. Eigen states of generators of rotations

We consider hermitian operators J_i, that satisfy the following commutation relations:[2]

$$[J_\alpha, J_\beta] = -\varepsilon_{\alpha\beta\gamma} J_\gamma \qquad [J^2, J_\alpha] = 0 \ .$$

[2]Take the representations (with different coefficients) from corollary D.1.18 for example.

Furthermore, we define the following operators

$$J^2 := \sum_{i=1}^{3} J_i^2 \qquad J^\pm := J_1 \pm iJ_2 \qquad \Rightarrow \quad (J^+)^\dagger = J^-$$

From the commutator relations for the J_i, the following equations can be shown:

$$[J_3, J^\pm] = \pm J^\pm \ , \qquad [J^+, J^-] = 2J_3 \ ,$$

$$J^2 = J^+J^- + J_3^2 - J_3 = J^-J^+ + J_3^2 + J_3 \ .$$

Theorem B.1.6 states, that J^2 and J_3 have a joint eigen basis $\{|jm\rangle\}$.

Remark B.3.1.
One could also choose the pairs J^2, J_1 or J^2, J_2. The choice of J^2, J_3 is only a convention.

In regard of the result, the notation $\{|jm\rangle\}$ was chosen, as is the notation for the eigen values:

$$J^2|jm\rangle = j(j+1)|jm\rangle \qquad J_3|jm\rangle = m|jm\rangle \ .$$

Theorem B.3.2.
For the eigen value numbers j and m of J^2 and J_3, the following restrictions apply:

$$j \in \mathbb{N}_0 \cup \tfrac{1}{2}\mathbb{N} \ ,$$
$$m \in \{-j, -j+1, ..., j-1, j\} \ .$$

Proof B.3.3.
The reason for the notation of $J^\pm$ as well as their meaning can be found from the commutator relations $[J_3, J^\pm] = \pm J^\pm$:

$$J_3J^\pm|jm\rangle = (J^\pm J_3 \pm J^\pm)|jm\rangle = (m \pm 1)J^\pm|jm\rangle \ .$$

Put differently, the operators J^+ and J^- map an eigenstate of J_3 to an eigen state of the eigen value increased and decreased by 1. Thus we may write

$$J^\pm|jm\rangle = \sqrt{j(j+1) - m(m \pm 1)}\ |j(m \pm 1)\rangle \ .$$

To see, that the coefficient $c_\pm = \sqrt{j(j+1) - m(m \pm 1)}$ is the right one, up to a phase $e^{i\varphi}$, we calculate

$$\begin{aligned} |c_+|^2 &= \langle J^+\ jm|J^+\ jm\rangle = \langle jm|J^-J^+|jm\rangle = \langle jm|J^2 - J_3^2 - J_3|jm\rangle \\ &= j(j+1) - m^2 - m = j(j+1) - m(m+1) \ , \end{aligned}$$

and similarly for $|c_-|^2$. The restriction for the eigen value numbers j and m is a consequence of the positive definiteness of the hermitian scalar product:

$$0 \leq \langle J_1\ jm|J_1\ jm\rangle + \langle J_2\ jm|J_2\ jm\rangle = \langle jm|J_1^2|jm\rangle + \langle jm|J_2^2|jm\rangle$$

$$= \langle jm|J_1^2 + J_2^2|jm\rangle = \langle jm|J^2 - J_3^2|jm\rangle = j(j+1) - m^2$$

$$\Leftrightarrow \quad |m| \leq \sqrt{j(j+1)}$$

Hence there is a minimal $m = \underline{m}$ and a maximal $m = \widetilde{m}$. Taking into account, the action of $J^\pm$ on $|jm\rangle$, and that $J^\pm|jm\rangle$ has to be an eigen state of J_3 to the eigen value $m \pm 1$, we need to demand:

$$\begin{aligned} J^+|j\widetilde{m}\rangle = 0 \quad &\Rightarrow \quad j(j+1) = \widetilde{m}(\widetilde{m}+1) \\ J^-|j\underline{m}\rangle = 0 \quad &\Rightarrow \quad j(j+1) = \underline{m}(\underline{m}-1) \ . \end{aligned}$$

Rewriting the system of equations yields:

$$\begin{aligned} &\widetilde{m} \geq \underline{m} \\ &\widetilde{m}(\widetilde{m}+1) = \underline{m}(\underline{m}-1) \ . \end{aligned}$$

It follows, that $\widetilde{m} = -\underline{m}$ and $j = \widetilde{m} = -\underline{m}$. Since $J^\pm$ increases/decreases in unit steps, j as to be in $\mathbb{N}_0 \cup \frac{1}{2}\mathbb{N}$. □

Remark B.3.4.
For the physical angular momentum, additional coefficients $\hbar$ are needed. However, the statement of the last theorem remains unchanged.

B.3.2. Addition of angular momentum and Clebsch-Gordon coefficients

We consider a system that allows for two angular momenta/spins. Again, ignoring the proper units, the total angular momentum operators are

$$\mathbf{J} = \vec{J_1} + \vec{J_2} \qquad J_z = J_{1z} + J_{2z} \quad ... \quad J^\pm = J_x \pm iJ_y \ ,$$

where J_i denotes the i-th particle: $J_{iz}|j_i, m_i\rangle = m_i|j_i, m_i\rangle$. As explained in subsection 2.1.1, the system is described by the tensor space of both Hilbert spaces. Thus, a tensor basis is:

$$|j_1, m_1 \ ; j_2, m_2\rangle = |j_1, m_1\rangle \otimes |j_2, m_2\rangle \ .$$

We are now looking for a joint eigen basis of J^2, J_z, J_i^2:

$$|j, m\rangle \coloneqq |j, j_1, j_2, m\rangle \ .$$

The completeness relation $\mathbb{1} = \sum_{j_i, m_i} |j_1, m_1; j_2, m_2\rangle\langle j_1, m_1; j_2, m_2|$ allows to write

$$|j, m\rangle = \sum_{\substack{\tilde{j}_1, \tilde{j}_2 \\ m_1, m_2}} |\tilde{j}_1, m_1; \tilde{j}_2, m_2\rangle\langle \tilde{j}_1, m_1; \tilde{j}_2, m_2|j, m\rangle \ .$$

Not all coefficients $\langle \tilde{j}_1, m_1; \tilde{j}_2, m_2|j, m\rangle$ are non-zero. By definition, $|j, m\rangle$ as well as $|\tilde{j}_1, m_1; \tilde{j}_2, m_2\rangle$ are eigen states of J_i^2:

$$\tilde{j}_i(\tilde{j}_i + 1)\langle \tilde{j}_1, m_1; \tilde{j}_2, m_2|j, m\rangle = \langle \tilde{j}_1, m_1; \tilde{j}_2, m_2|J_i^2|j, m\rangle$$

$$= j_i(j_i+1)\langle \tilde{j}_1, m_1; \tilde{j}_2, m_2 | j, m\rangle \ .$$

Hence, if $j_i \neq \tilde{j}_i$, it must hold that $\langle \tilde{j}_1, m_1; \tilde{j}_2, m_2 | j, m\rangle = 0$.

Furthermore, $J_z = J_{1z} + J_{2z}$ causes the addition of eigen values $m = m_1 + m_2$, such that:

$$\begin{aligned} m\langle \tilde{j}_1, m_1; \tilde{j}_2, m_2 | j, m\rangle &= \langle \tilde{j}_1, m_1; \tilde{j}_2, m_2 | J_z | j, m\rangle \\ &= \langle \tilde{j}_1, m_1; \tilde{j}_2, m_2 | J_{1z} + J_{2z} | j, m\rangle \\ &= (m_1+m_2)\langle \tilde{j}_1, m_1; \tilde{j}_2, m_2 | j, m\rangle \ . \end{aligned}$$

If $m_1 + m_2 \neq m$, then it must hold that $\langle \tilde{j}_1, m_1; \tilde{j}_2, m_2 | j, m\rangle = 0$. Finally, taking only the non-zero coefficients into account, we can write:

$$|j, m\rangle = \sum_{\substack{m_1, m_2 \\ m_1+m_2=m}} |j_1, m_1; j_2, m_2\rangle \langle j_1, m_1; j_2, m_2 | j, m\rangle \ .$$

Definition B.3.5.
The coefficients $\langle j_1, m_1; j_2, m_2 | j, m\rangle$ are called **Clebsch-Gordon coefficients**.

So far, the Clebsch-Gordon coefficients are characterized by the condition, that $\{|j, m\rangle\}$ has to be a joint eigen basis of the operators J^2, J_z, J_i^2. It remains to find coefficients, such that the condition is met. In fact, there do exist algorithms to find such coefficients.

C

Tensors and Index-Notation

Tensors are a vital concept in physics. This chapter tries to introduce the underlying algebraic concept of tensor products following [RW05] and [HO07]. As result, the universal property is the starting point, allowing to prove existence and the well known properties for calculations. We conclude this chapter with an introduction to Ricci-calculus, as given in [Jän05], observing the invariant isomorphisms behind index manipulations.

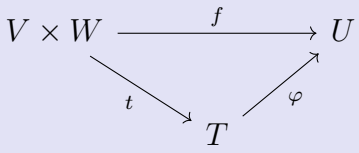
C.1. Tensor product

Usually (at least in the physical literature) tensors are defined by transformation behavior or calculational properties. Though this approach delivers a ready introduction for calculations, the concepts remain non-transparent. The most transparent way, yet sadly also the most abstract way, is the definition by universal properties commonly used in abstract algebra. However, the fundamental behavior, the transformation behavior, follows as direct corollary. An important structure left behind in the approach of transformation behavior, is the tensor product, that will be the first object to investigate here.

C.1.1. Existence and uniqueness

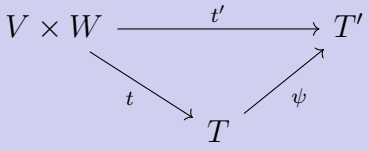
The following section is rather technical and can be skipped, if one is only interested in the behavior of tensors.

Definition C.1.1 (Universal property)**.**
Let V and W be $\mathbb{K}$-vector spaces. Also let (T, t) be a tupel, consisting of a vector space T and a bilinear map $t\colon V \times W \to T$. The tuple is called **tensor product** if the following universal property is fulfilled:
Let U be another vector space and $f\colon V \times W \to U$ be a linear map. Then there exists a linear map $\varphi_f\colon T \to U$, such that the following diagram commutes:



Theorem C.1.2 (Existence and uniqueness)**.**
For any two $\mathbb{K}$- vector spaces V and W, there exists always a tensor product (T, t). This tensor product is unique up to isomorphy. That is, if (T', t') is a second tensor product, then there exists a defined isomorphism $\Psi\colon T \to T'$ such

that the following diagram commutes:



Proof C.1.3.
Preparations: Let $\mathbb{K}^M$ denote the set of maps from M in the field $\mathbb{K}$. The set $\mathbb{K}^M$ is a vector space with point wise addition. A special class of maps in this vector space are *Kronecker-deltas*

$$\delta_m \colon M \longrightarrow \mathbb{K} \ , \quad x \longmapsto \delta_m(x) \coloneqq \begin{cases} 1 & , x = m \\ 0 & , \text{sonst} \end{cases} \ .$$

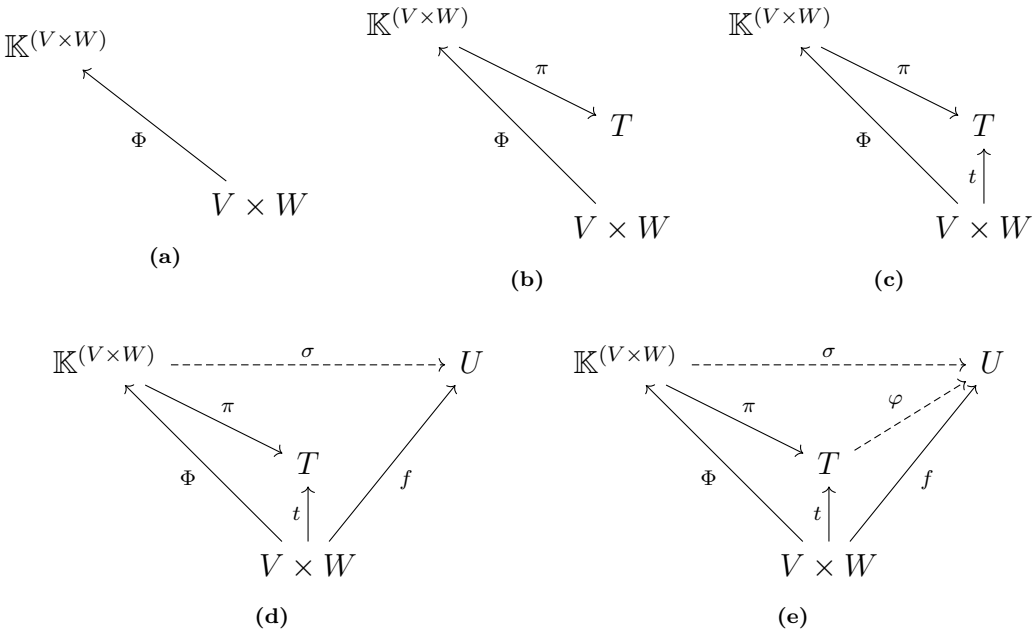
This gives rise to an embedding Φ of M in $\mathbb{K}^M$:

$$\Phi \colon M \hookrightarrow \mathbb{K}^M \ , \quad m \mapsto \delta_m \ .$$

An important subspace of $\mathbb{K}^M$ for this proof is the space of maps with $f(x) = 0$ for all but finitely many $x \in M$. This subspace will be denoted by $\mathbb{K}^{(M)}$. Every element of this space can be written as linear combination of finitely many deltas δ_m. By definition the deltas are linear independent and thus forming a basis of $\mathbb{K}^{(M)}$:

$$\mathbb{K}^{(M)} = \operatorname{span}_{\mathbb{K}}(im(\Phi)) \ .$$

Existence:
The existence follows from inspecting a commutative diagram:



(a) As we have seen, there is an embedding, that is, an injective map $\Phi\colon V \times W \to \mathbb{K}^{(V\times W)}$.

(b) We define the subspace $X \subset \mathbb{K}^{(V\times W)}$. Let $v, v' \in V$, $w, w' \in W$ and $a \in \mathbb{K}$, then X shall be defined as linear span of the following elements

$$\delta_{(v+v',w)} - \delta_{(v,w)} - \delta_{(v',w)} \ , \quad \delta_{(v,w+w')} - \delta_{(v,w)} - \delta_{(v,w')} \ ,$$

$$\delta_{(av,w)} - \delta_{a(v,w)} \quad \text{and} \quad \delta_{(v,aw)} - \delta_{a(v,w)} \ .$$

The space T can now be defined as quotient $T \coloneqq \mathbb{K}^{(V\times W)}/X$. Hence T is a space of equivalence classes with the following equivalence relation:

$$h \sim h' \quad \Leftrightarrow \quad \exists\, x \in X\colon x' = h + x \ .$$

Let $\pi\colon \mathbb{K}^{(V\times W)} \to T$ be the canonical projection, i.e. the surjective map assigning every $h \in \mathbb{K}^{(V\times W)}$ its equivalence class $\pi(h) = [h] \in \mathbb{K}^{(V\times W)}/X$.

(c) The map $t\colon V \times W \to T$ will be defined by $t = \pi \circ \Phi$. Due to the choice of X the map is bilinear:

$$t((v+v',w) - (v,w) - (v',w)) = \pi(\delta_{(v+v',w)} - \delta_{(v,w)} - \delta_{(v',w)}) = [0]$$

$$\Rightarrow \quad [\delta_{(v+v',w)}] = [\delta_{(v,w)}] + [\delta_{(v',w)}] \quad \Leftrightarrow \quad t((v+v',w)) = t((v,w)) + t((v',w)) \ .$$

The remaining properties can be shown similarly.

(d) It remains to show, that (T, t) satisfies the universal property. So let U be a $\mathbb{K}$-vector space and $f\colon V \times W \to U$ a bilinear map. The image of Φ defines a basis of $\mathbb{K}^{(V\times W)}$. Define the map σ by

$$\sigma(\Phi(v,w)) = f(v,w) \ .$$

By linear completion σ is a linear map $\mathbb{K}^{(V\times W)} \to U$. From the bilinearity of f follows the bilinearity of σ:

$$\sigma(\Phi(v+v',w)) = f((v+v',w)) = f(v,w) + f(v',w) = \sigma(\Phi(v,w)) + \sigma(\Phi(v',w)) \ .$$

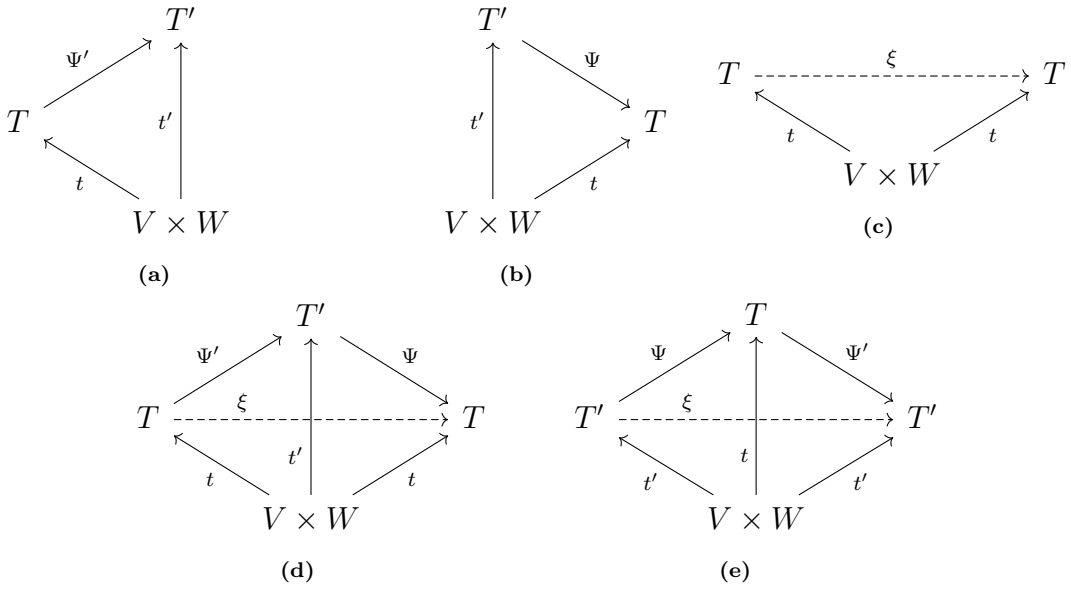
Hence $U \subseteq \ker(\sigma)$.

(e) The fundamental theorem on homomorphisms states, that there exists a unique linear map φ, completely defined by σ, such that the upper triangle of the diagram commutes. Since σ is uniquely defined by f, so is φ, proving the universal property.

Uniqueness:
To prove uniqueness we use a second tensor space T' for U:

(a) Since we have already proven the universal property, we know that there is a unique map $\Psi'\colon T \to T'$, defined by t'.

(b) Similarly there is a unique map $\Psi\colon T' \to T$ defined by t.



(c) Also there is a unique linear map $\xi \in \mathrm{End}(T)$, such that

$$\xi \circ t = t$$

holds. Furthermore, $Id_T \circ t = t$ holds. Since ξ was unique, $\xi = Id_T$.

(d) Combining the diagrams, which commute by definition, shows that the whole diagram does, too. Hence:

$$Id_T = \xi = \Psi' \circ \Psi \ .$$

(e) Similarly it follows that $Id_{T'} = \Psi \circ \Psi'$. Due to the uniqueness of Ψ, being determined by t and the uniqueness of Ψ', being determined by t', theses maps coincide in both diagrams. Thus finally $\Psi' = \Psi^{-1}$, i.e. it is an isomorphism.

□

C.1.2. Tensors

The tensor product is unique up to isomorphy, so it is common to speak about the tensor product. The usual notation for (T, t) is $(V \otimes W, \otimes)$. As long as the field $\mathbb{K}$ is understood, one can write $\otimes$, otherwise one needs to specify the field, e.g. $\otimes_{\mathbb{K}}$. This is important, since $\otimes$ is bilinear only with respect to $\mathbb{K}$:

$$(v + v') \otimes w = v \otimes w + v' \otimes w \ ,$$

$$v \otimes (w + w') = v \otimes w + v \otimes v' \ ,$$

$$(\alpha v) \otimes w = \alpha(v \otimes w) = v \otimes (\alpha w) \ .$$

Definition C.1.4.
Tensors of the form $v \otimes w \in V \otimes W$ with $v \in V$ and $w \in W$ are called **pure tensors**

By the construction of $V \otimes W$ it follows that every tensor can be written as sum of pure tensors.

Theorem C.1.5.
Let $\{v_i\}_{i\in I}$ be a basis of V and $\{w_j\}_{j\in J}$ a basis of W for finite dimensional vector spaces. Then $\{v_i \otimes v_j\}_{i\in I, j\in J}$ is a basis of $V \otimes W$ and it follows that:

$$\dim(V \otimes W) = \dim(V) \cdot \dim(W) \ .$$

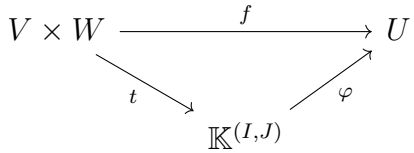
Proof C.1.6.
We consider the following map:

$$t' \colon V \times W \longrightarrow \mathbb{K}^{(I,J)} \ , \qquad \left(\sum_{i\in I} x_i v_i, \sum_{j\in J} y_j w_j\right) \longmapsto x_i y_j \cdot \delta_{(i,j)} \ .$$

Here I and J are the index sets of the bases of V and W. The map t' is bilinear and maps elements (v_i, w_j) to the basis elements $\delta_{(i,j)}$ of $\mathbb{K}^{(I,J)}$. By linear completion there is a unique linear map $\varphi \colon \mathbb{K}^{(I,J)} \to U$ for every bilinear map $f \colon V \times W \to U$, defined by

$$f(v, w) = \varphi(t'(v, w)) \qquad \forall\ v \in V,\ w \in W \ .$$

Thus the following diagram commutes:



Hence the tuple $(\mathbb{K}^{(I,J)}, t')$ satisfies the universal property. With theorem C.1.2 we find $\mathbb{K}^{(I,J)} \cong V \otimes W$. The isomorphism Ψ between those vector spaces has the property $\Psi \circ t' = \otimes$, mapping $t'(v_i, w_j) = \delta_{(i,j)}$ to $v_i \otimes w_j$. Since $\{\delta_{(i,j)}\}$ is a basis of $\mathbb{K}^{(I,J)}$, so is $\{v_i \otimes w_j\}$ a basis of $V \otimes W$. □

With the universal property we can prove the following isomorphies:

Lemma C.1.7 (Ismorphisms of tensor spaces)**.**
Let V, W and U be $\mathbb{K}$-vector spaces, then the following isomorphies are unique for the stated conditions:

(i) $V \otimes W \cong W \otimes V, \quad v \otimes w \mapsto w \otimes v.$

(ii) $(U \otimes V) \otimes W \simeq U \otimes (V \otimes W), \quad (u \otimes v) \otimes w \mapsto u \otimes (v \otimes w).$

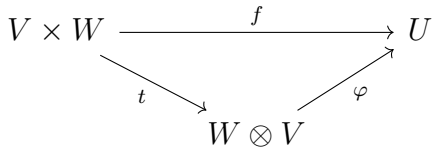
(iii) $(U \bigoplus V) \otimes W \simeq (U \otimes W) \bigoplus (V \otimes W) \quad (u, v) \otimes w \mapsto (u \otimes w, v \otimes w)$.

(iv) $\mathbb{K} \otimes_{\mathbb{K}} V \cong V, \quad a \otimes v \mapsto a \cdot v$.

Proof C.1.8 (Œ for (i))**.**
Let $t\colon V \times W \to W \otimes V$ be defined by $t(v, w) = w \otimes v$. Every bilinear map $f\colon V \times W \to U$ determines $\varphi\colon W \times V \to U$ uniquely by

$$\varphi(w \otimes v) = f(v, w) \ ,$$

such that the following diagram commutes:



Thus $(W \otimes V, t)$ satisfies the universal property and is isomorphic to $V \otimes W$ due to theorem C.1.2. The remaining isomorphisms can be proven similarly.[1] □

The last isomorphism (iv) is only valid, if the field used as vector space is the same as the field used to define the tensor product. If $\mathbb{K}' \supset \mathbb{K}$ is a field containing $\mathbb{K}$ as subfield (e.g. $\mathbb{C}$ and $\mathbb{R}$), then statement (iv) fails: $\mathbb{K}' \otimes_{\mathbb{K}} V \ncong V$. Yet $\mathbb{K}' \otimes_{\mathbb{K}} V$ becomes a $\mathbb{K}'$-vector space. That is, the scalar range of V is extended to $\mathbb{K}'$, by

$$a' \cdot (b' \otimes v) = (a' \cdot b') \otimes v \ .$$

Definition C.1.9.
Let $\mathbb{K}$ be a subfield of $\mathbb{C}$, i.e. $\mathbb{K} = \mathbb{Q}$ or $\mathbb{R}$, and let V be a $\mathbb{K}$-vector space. The tensor product $V^{\mathbb{C}} := \mathbb{C} \otimes_{\mathbb{K}} V$ is called **complexification** of V.

It can be shown, that if $\{v_j\}_{j \in J}$ is a basis of V, so is $\{1 \oplus_{\mathbb{K}} v_j\}_{j \in J}$ of $V^{\mathbb{K}'}$.

Lemma C.1.10.
Let V_1, V_2, W_1, W_2 be vector spaces, and $\varphi_1\colon V_1 \to W_1$ as well as $\varphi_2\colon V_2 \to W_2$ be linear maps. Then there is a unique linear map

$$\varphi_1 \otimes \varphi_2\colon V_1 \otimes V_2 \longrightarrow W_1 \otimes W_2 \ , \qquad (\varphi_1 \otimes \varphi_2)(v_1 \otimes v_2) = \varphi_1(v_1) \otimes \varphi_2(v)_2 \ .$$

This tensor product of linear maps has the following properties:

1. $Id_{V_1} \otimes Id_{V_2} = Id_{V_1 \otimes V_2}$.
2. *For two additional linear maps $\varphi_1'\colon V_1 \to W_1$ and $\varphi_2'\colon V_2 \to W_2$, it holds that:*
$$(\varphi_1 \otimes \varphi_2) \circ (\varphi_1' \otimes \varphi_2') = (\varphi_1 \circ \varphi_1') \otimes (\varphi_2 \circ \varphi_2') \ .$$

[1] For example see [RW05, p. 4-5] for (iv).

3. *The map*

$$\begin{aligned}\mathrm{Hom}(V_1, W_1) \times \mathrm{Hom}(V_2, W_2) &\hookrightarrow \mathrm{Hom}(V_1 \otimes V_2, W_1 \otimes W_2)\\ (\varphi_1, \varphi_2) &\mapsto \varphi_1 \otimes \varphi_2\end{aligned}$$

is bilinear and injective. For finite-dimensional vector spaces it is an isomorphism.

From the definition of the algebraic dual space $V^* = \mathrm{Hom}(V, \mathbb{K})$ and the isomorphy $\mathbb{K} \otimes_{\mathbb{K}} V \cong V$ the injection

$$V_1^* \otimes V_2^* \hookrightarrow (V_1 \otimes V_2)^* \ ,$$

follows as direct result from the previous lemma. Accordingly it is an isomorphism for finite-dimensional vector spaces.

Lemma C.1.11.
There is an embedding $V_2 \otimes V_1^*$ into $\mathrm{Hom}(V_1, V_2)$, defined by the following injective linear map

$$V_2 \otimes V_1^* \hookrightarrow \mathrm{Hom}(V_1, V_2) \ , \quad v_2 \otimes \vartheta_1 \mapsto \ell_{v_2, \vartheta_1} \ ,$$

where $\ell_{v_2,\vartheta_1}(v) = \vartheta_1(v) \cdot v_2$. This map can be extended linearly for $V_2 \otimes V_1^*$.

As before, the embedding becomes an isomorphy in the finite-dimensional case.

Corollary C.1.12.
Let $\{e_i\}_i$ be a basis of a finite-dimensional vector space V and $\{\vartheta_j\}_j$ the dual basis. Then, every linear operator $L \in \mathrm{Hom}(V, W)$ can be written as tensor from $W \otimes V^*$:

$$L = \sum_i (Le_i) \otimes \vartheta_i \ .$$

Although the tensor product is defined for infinite-dimensional vector spaces, too, the tensor product of two Hilbert spaces does not need to be one as well. The missing property is completeness here.

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two Hilbert spaces. One can define a scalar product on $\mathcal{H}_1 \otimes \mathcal{H}_2$ by

$$\langle v_1 \otimes v_2 | w_1 \otimes v_2 \rangle = \langle v_1 | w_1 \rangle \cdot \langle v_2 | w_2 \rangle \ .$$

To obtain a Hilbert space, one can take the metric completion with respect to that scalar product.

We have spoken about bases so far, even in the infinite-demensional case. A basis allows to linear combine every element of the space with finitely many basis elements.[2] Theorem C.1.5 can be extended to Hilbert-bases.

[2] Contrasting Hilbert-bases, where the sum is infinite. In infinite spaces, a true basis may even become uncountable.

C.2. Ricci calculus

In Ricci calculus, tensors are characterized by their coefficients (becoming a list of numbers). There are three important principles we need to understand:

1) Components define objects
2) Position of indices determines transformation behavior.
3) Summation convention.

C.2.1. Co-and contravarianz

Let V be a vector space, $\{e_i\}_{i=1,\dots,n}$ a basis and V^* be the dual space with dual basis $\{\vartheta^i\}_{i=1,\dots,n}$. A vector $v \in V$ is called **contravariant** and is described by coefficients with upper indices:

$$v = \sum_{i=1}^{n} v^i e_i \equiv v^i e_i \ .$$

A dual vector $\varphi \in V^*$ is called **covariant** and is described by coefficients with lower indices

$$\varphi = \varphi_i \vartheta^i \ .$$

In the last equation we have already used the summation convention. Over same indices, one upper and one lower, will always be summed (without having to write the summation symbol).

Changing the basis $e_i \to \tilde{e}_i$ does not change the element v, but its components:

$$v = v^i e_i = \tilde{v}^k \tilde{e}_k \ .$$

By definition there are coefficients, such that $e_i = A^j{}_i \tilde{e}_j$ can be written. Plugging in yields the connection between v^i and $\tilde{v}^k$:

$$v = \tilde{v}^k \tilde{e}_k = v^i e_i = v^i A^j{}_i \tilde{e}_j \qquad \Rightarrow \qquad \tilde{v}^k = A^k{}_i v^i \ .$$

For the coefficients of the matrix A, we have already used the Ricci convention. Still, the first index describes the row and the second index the columns, independent if it is an upper or lower index.

Remark C.2.1 (Composition of linear maps)**.**
Let $A, B \in \mathrm{End}(V)$ for a finite dimensional vector space V. Due to the isomorphy $\mathrm{End}(V) \simeq V \otimes V^*$, these maps can be written as tensors:

$$A = A e_i \otimes \vartheta^i = A^j{}_i e_j \otimes \vartheta^i \quad \text{and} \quad B = B^k{}_\ell e_k \otimes \vartheta^\ell \ .$$

Evaluating $(A \circ B)(v)$ for an arbitrary vector $v \in V$ yields

$$\begin{aligned}(A \circ B)(v) &= A^j{}_i e_j \cdot \vartheta^i \left(B^k{}_\ell e_k \cdot \vartheta^\ell(v) \right) = A^j{}_i B^k{}_\ell \vartheta^\ell(v) \vartheta^i(e_k) \cdot e_j \\ &= A^j{}_i B^k{}_\ell \vartheta^\ell(v) \delta^i_k \cdot e_j = A^j{}_i B^i{}_\ell \vartheta^\ell(v) \cdot e_j \\ &= \left(\left(A^j{}_i B^i{}_\ell \right) e_j \otimes \vartheta^\ell \right)(v) \ .\end{aligned}$$

From the last equality we can read off the coefficient behavior under composition:

$$\boxed{(A \circ B)^j{}_\ell = A^j{}_i B^i{}_\ell}$$

If A is a basis change matrix, there is an inverse A^{-1}. From the above remark we know, that this can be expressed by $A^i{}_j (A^{-1})^j{}_k = \delta^i_k$. Dual vectors are linear and defined by $\vartheta^i(e_j) = \delta^i_j$ and respectively $\tilde{\vartheta}^i(\tilde{e}_j) = \delta^i_j$, thus we find:

$$\delta^i_j = \vartheta_i(e_j) = \vartheta_i(A^k{}_j \tilde{e}_k) = A^k{}_j \vartheta_i(\tilde{e}_k) \ .$$

Since the dual vector space is also a vector space, there are coefficients such that $\vartheta^i = M^i{}_\ell \tilde{\vartheta}^\ell$. Plugging in results in the transformation behavior of covectors:

$$\delta^i_j = A^k{}_j \vartheta^i(\tilde{e}_k) = A^k{}_j M^i{}_\ell \tilde{\vartheta}^\ell(\tilde{e}_k) = A^k{}_j M^i{}_\ell \delta^k_\ell = A^k{}_j M^i{}_k = M^i{}_k A^k{}_j \quad \Rightarrow \quad M = A^{-1} \ .$$

$$\Rightarrow \qquad \varphi = \tilde{\varphi}_j \tilde{\vartheta}^j = \varphi_i \vartheta^i = \varphi_i (A^{-1})^i{}_k \tilde{\vartheta}^k \qquad \Rightarrow \qquad \tilde{\varphi}_j = (A^{-1})^i{}_j \varphi_i \ .$$

By definition, basis change matrices are orthogonal/unitary. That is $A^{-1} = A^\dagger$ / $A^{-1} = A^T$. Summing up our findings:

	coefficients	**transformation basis vectors**
contravariant	$\tilde{v}^k = A^k{}_i v^i$	$\tilde{e}_k = (A^T)^i{}_k e_i$
covariant	$\tilde{\varphi}_k = (A^T)^i{}_k \varphi_i$	$\tilde{\vartheta}^k = A^k{}_i \vartheta^i$

C.2.2. Tensors in Ricci calculus

After we have seen the foundations of Ricci calculus we can use this formulation on tensors:

Definition C.2.2.
A tensor, consisting of r vectors and s covectors

$$T \in \underbrace{V \otimes ... \otimes V}_{r \text{ times}} \otimes V^* \otimes \underset{s \text{ times}}{... \otimes V^*}$$

is called **tensor of type (r, s)**. The number $r + s$ is called the **rank**, also for a general order of vectors and covectors.

A tensor of type (r, s) can be expanded as follows:

$$T = T^{i_1 ... i_r}{}_{j_1 ... j_s} \, e_{i_1} \otimes ... \otimes e_{i_r} \otimes \vartheta^{j_1} \otimes ... \otimes \vartheta^{j_s} \ .$$

In Ricci calculus one agrees upon the following identification:

$$T = T^{i_1 ... i_r}{}_{j_1 ... j_s} \ .$$

A change of basis results in the following transformation behavior:

$$\tilde{T}^{n_1 ... n_r}{}_{m_1 ... m_s} = A^{n_1}{}_{i_1} ... A^{n_r}{}_{i_r} (A^{-1})^{m_1}{}_{j_1} ... (A^{-1})^{m_s}{}_{j_s} \, T^{i_1 ... i_r}{}_{j_1 ... j_s}$$

This behavior is used to define tensors in the physical literature.

C.2.3. Raising and lowering indices

In case of a Riemannian manifold there is a fourth principle, induced by the Riemannian metric:

4) Raising and lowering indices

To understand the invariant meaning behind these manipulations, instead of just defining them, it is best to use the coordinate free formulation first.

Let g be a scalar product (or Riemannian metric). Then there is an isomorphism

$$I_1 \colon \text{vector} \longrightarrow \text{covector} \qquad v \longmapsto I_1(v) = g(v, \cdot) \ .$$

Remark C.2.3.
In the literature, the isomorphism I_1 and its inverse I_1^{-1} are called **flat-** and **sharp isomorphism** respectively. The usual notation is

$$I_1(v) = v^\flat \qquad \text{and} \qquad I_1^{-1}(\omega) = \omega^\# \ .$$

These isomorphisms can be applied to individual parts of the tensor, still defining an isomorphism between tensor spaces. For example, a $(1,1)$-tensor becomes a $(0,2)$-tensor if $I_1 \otimes \mathbb{1}$ is applied, and a $(2,0)$-tensor, if $\mathbb{1} \otimes I_1^{-1}$ is applied.

Let ∂_μ be a tangent basis and dx^ν be the dual basis of the cotangent space. Defining the coefficients of the Riemannian metric by $g_{\mu\nu} = g(\partial_\mu, \partial_\nu)$, there is an inverse matrix (list of numbers) $g^{\mu\nu}$. By definition of scalar products the matrices are symmetric: $g_{\mu\nu} = g_{\nu\mu}$ and hence $g^{\mu\nu} = g^{\nu\mu}$. For the isomorphisms I_1 and I_1^{-1} it follows that:

$$I_1(\partial_\mu) = g(\partial_\mu, \cdot) = g_{\mu\nu} dx^\nu \qquad \text{and thus}[3] \qquad I_1^{-1}(dx^\mu) = g^{\mu\nu} \partial_\nu \ .$$

The coefficients transform as follows:

$$I_1(v^\mu \partial_\mu) = g_{\mu\nu} v^\mu dx^\nu =: v_\nu dx^\nu \qquad \text{and} \qquad I_1^{-1}(u_\mu dx^\mu) = g^{\mu\nu} u_\mu \partial_\nu =: u^\nu \partial_\nu \ .$$

Remark C.2.4.
A contravariant vector v^μ becomes a covariant vector v_ν by lowering the index:

$$v_\nu = g_{\mu\nu} v^\mu \ .$$

Raising an index on the other hand, transforms a covector into a vector:

$$u^\nu = g^{\mu\nu} u_\mu \ .$$

The raising and lowering can be applied for indices of tensors separately:

$$g_{\mu\nu} A^{\cdots\mu\cdots}{}_{\cdots} = A^{\cdots}{}_{\mu}{}^{\cdots}{}_{\cdots} \ .$$

[3] $\partial_\mu = I_1^{-1}(g_{\mu\nu} dx^\mu) = g_{\mu\nu} I_1^{-1}(dx^\mu) \quad \Rightarrow \quad g^{\mu\nu} g_{\mu\nu} I_1^{-1}(dx^\mu) = I_1^{-1}(dx^\mu) = g^{\mu\nu} \partial_\mu$.

Remark C.2.5.
The scalar product of two vectors $u^\mu \partial_\mu$ and $v^\nu \partial_\nu$ can be written as composition:

$$g(u^\mu \partial_\mu, v^\nu \partial_\nu) = u^\mu v^\nu g_{\mu\nu} = u_\nu v^\nu = v_\mu u^\mu \ .$$

D
Representation theory

In this chapter, the basics of representation theory, that can be found throughout [Zir13] and [Zir10], are summarized. The definitions and concepts of representations of groups and algebras as well as important algebras for quantum mechanics and their representations are introduced.

D.1. Overview of representation theory

Representations allow to map structures onto vector spaces to investigate properties in the well understood context of linear maps. In the following we will consider groups and algebras.

D.1.1. Representations of groups

Definition D.1.1.
Let G be a group and V be a $\mathbb{K}$ vector space. A group homomorphism $\mathcal{D} : G \to \mathrm{GL}(V)$ is called **representation** of G on V.

A **group homomorphism** is a map $H : G_1 \to G_2$ between two groups, that satisfies

$$H(g \circ k) = H(g) \circ H(k) \quad \forall\ g, k \in G_1\ .$$

A representation is the special case of G_2 being the general linear group of a vector space.

Definition D.1.2.
If V is a complex vector space with hermitian scalar product $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ and the representation satisfies

$$\langle \mathcal{D}(g)v, \mathcal{D}(g)w \rangle = \langle v, w \rangle \quad \forall\ v, w \in V \ \ \forall g \in G\ ,$$

the representation si called **unitary**.

Instead of using group homomorphisms, a group representation can be constructed using group actions on a vector space:

Definition D.1.3.
Let $(G, \circ)$ be a group and M a set. A **group action** of the group G on the set M is a binary operator $\diamond : G \times M \to M$ wit the following properties for all $g_1, g_2 \in G$ and $x \in M$:

$$(g_1 \circ g_2) \diamond x = g_1 \diamond (g_2 \diamond x)$$

$$e \diamond x = x \ ,$$

where e is the neutral element of the group.

Theorem D.1.4.
Let G be a group and V be a $\mathbb{K}$ vector space. A linear group action of G on V defines a group representation of G on V.

Proof D.1.5.
To be linear means, that the group action has the following property:

$$g \diamond (\alpha v + \beta w) = \alpha \cdot g \diamond v + \beta \cdot g \diamond w \qquad \forall v, w \in V \ \alpha, \beta \in \mathbb{K} \ .$$

Because of the linear group action, a map $\Phi \colon G \to \mathrm{End}(V)$ can be defined by:

$$g \mapsto \Phi_g \quad \text{mit} \quad \Phi_g(v) = g \diamond v \quad \forall g \in G \ \forall v \in V \ .$$

The linearity of Φ_g is a direct consequence of the linearity of the group action. Using the properties of group actions show that Φ_g is a group homomorphism:

$$\Phi_{g_1 \circ g_2} v = (g_1 \circ g_2) \diamond v = g_1 \diamond (g_2 \diamond v) = (\Phi_{g_1} \circ \Phi_{g_2}) v \ .$$

Let g^{-1} be the inverse element of $g \in G$, then:

$$\Phi_{g^{-1}} \circ \Phi_g = \Phi_{g^{-1} \circ g} = \Phi_e = Id_V = \ldots = \Phi_g \circ \Phi_{g^{-1}} \quad \Rightarrow \quad \Phi_{g^{-1}} = \Phi_g^{-1} \ .$$

Hence every Φ_g is invertible and thus $\Phi \colon G \to \mathrm{GL}(V)$. □

To close this subsection we will prove Schur's lemma (as can be found in [Zir10, section 4.2]). To do so we need further definitions:

Definition D.1.6.
A representation $\mathcal{D}$ of G on V is called **irreducible**, if the only G-invariant subspaces if V are $\{0\}$ and V itself. A subspace $U \subset V$ is called ***G*-invariant** , if

$$\mathcal{D}(g)U = U \quad \forall \ g \in G \ .$$

Definition D.1.7.
Two representations $(G, V_1, \mathcal{D}_1)$ and $(G, V_2, \mathcal{D}_2)$ are called **equivalent**, if there exists an isomorphism $\Phi \colon V_1 \to V_2$ such that $\Phi \circ \mathcal{D}_1 = \mathcal{D}_2 \circ \Phi$ for all $g \in G$.

Theorem D.1.8 (Schur's lemma).
Let $(G, V_1, \mathcal{D}_1)$ and $(G, V_2, \mathcal{D}_2)$ be two irreducible maps of G on a finite-

dimensional vector space. If there is a linear map $\Phi\colon V_1 \to V_2$ *that satisfies*

$$\Phi \circ \mathcal{D}_1(g) = \mathcal{D}_2(g) \circ \Phi \quad \forall\ g \in G\ ,$$

then either $\Phi \equiv 0$ *or* Φ *is invertible, i.e. an isomorphism.*

Proof D.1.9.
Let $v \in \ker(\Phi) \subset V_1$, then for all $g \in G$ it holds that:

$$0 = \Phi(v) \quad \Rightarrow \quad 0 = \mathcal{D}_2(g)\Phi(v) = \Phi(\mathcal{D}_1(g)v) \quad \Rightarrow \quad \mathcal{D}_1(g)v \in \ker(\Phi)\ .$$

Thus $\ker(\Phi)$ is a G-invariant subspace of V_1. By assumption it holds that $\ker(\Phi) = V_1$, which would lead to $\Phi \equiv 0$, or $\ker(\Phi) = \{0\}$.

Assuming $\ker(\Phi) = \{0\}$, then Φ is injective and thus also bijective by the rank-nullity theorem. □

Corollary D.1.10.
Let $(G, V, \mathcal{D})$ *be an irreducible representation on a finite-dimensional* $\mathbb{C}$ *vector space. Then, every endomorphism* $\Phi \in \mathrm{End}(V)$ *that satisfies*

$$\Phi \circ \mathcal{D}(g) = \mathcal{D}(g) \circ \Phi \qquad \forall\ g \in G$$

is proportional to the identity map, i.e. $\Phi \sim Id_V$.

Proof D.1.11.
In the complex number field the characteristic polynomial $\det(\Phi - x \cdot Id_v)$ has at least one zero (exactly $\dim(V)$ to be precise). Let λ be the zero, i.e. as eigen value of Φ.

The map $\Phi - \lambda Id_V$ satisfies

$$(\Phi - \lambda Id_V) \circ \mathcal{D}(g) = \mathcal{D}(g) \circ (\Phi - \lambda Id_V)\ ,$$

since Id_V commutes with every $\mathcal{D}(g)$. By definition $(\Phi - \lambda Id_V)$ is not invertible since the kernel is not zero. Using Shur's lemma this means

$$(\Phi - \lambda Id_V) = 0 \quad \Rightarrow \quad \Phi = \lambda Id_V \sim Id_V\ .$$

□

D.1.2. Representations of algebras

To every Lie group there is a Lie algebra. Thus we are also interested in representations of algebras.

Definition D.1.12.
Let V be a $\mathbb{K}$ vector space and $\circ\colon V \times V \to V$ a bilinear operation, then the pair $(V, \circ)$ is called **$\mathbb{K}$-algebra**. If the operation $\circ$ is associative the algebra is called **associative algebra**

The distinction between associative and non-associative algebras is an important one. Indeed, an important class of algebras we have already encountered, the Lie algebras, are non-associative.

Definition D.1.13.
Let $\mathcal{A}$ be a $\mathbb{K}$-algebra. An algebra homomorphism $\mathcal{D}\colon \mathcal{A} \to \mathrm{End}(V)$ is called **representation of the $\mathbb{K}$-algebra** on the vector space. In the case of a Lie algebra a **Lie algebra representation** is a Lie algebra homomorphism $\mathcal{D}\colon \mathcal{A} \to \mathfrak{gl}(V) \cong (\mathrm{End}(V), [\bullet, \bullet])$.

An **algebra homomorphism** is a linear map $H\colon \mathcal{A} \to \mathcal{A}'$ between two algebras, such that

$$H(x \circ y) = H(x) \circ H(y) \quad \forall\, x, y \in \mathcal{A}\ .$$

In the case of Lie algebras the condition is

$$H([x, y]) = [H(x), H(y)] \quad \forall\, x, y \in \mathcal{A}\ ,$$

of course.

Theorem D.1.14.
Let $\mathcal{D}\colon G \to \mathrm{GL}(V)$ be a smooth representation of a Lie group. The differential in the neutral element is a representation $D_e\mathcal{D}\colon \mathfrak{g} \to \mathfrak{gl}(V)$ of the associate Lie algebra.

Proof D.1.15.
It can be shown that $\mathrm{GL}(V)$ is a Lie group and $\mathfrak{gl}(V)$ its Lie algebra. The differential is by definition a linear map between Lie algebras (see lemma B.1.16):

$$D_e\mathcal{D}\colon T_eG \cong \mathfrak{g} \longrightarrow T_{\mathcal{D}(e)}\mathrm{GL}(V) \cong \mathfrak{gl}(V)\ .$$

Lemma B.1.10 shows that $D_e\mathcal{D}([X, Y]) = [D_e\mathcal{D}(X), D_e\mathcal{D}(Y)]$ holds. □

Lemma and example D.1.16.
A unitary representation of the rotation group $\mathrm{SO}(3)$ on the Hilbert space $L^2(\mathbb{R}^3)$ is given by

$$(\mathcal{D}(g)f)(x) := f(g^{-1}x)\ \forall x \in \mathbb{R}^3\ .$$

Hence $\mathcal{D}_*$ defined by

$$\mathcal{D}_*(X) = D_e\mathcal{D}(X) = \left.\frac{d}{d\varphi}\right|_{\varphi=0} \mathcal{D}(e^{\varphi X})$$

is a representation of $\mathfrak{so}(3)$.

Proof D.1.17.
To see more easily that $\mathcal{D}$ is a homomorphism it is helpful to notice that $g \in \mathrm{SO}(3)$ is a diffeomorphism $\mathbb{R}^3 \to \mathbb{R}^3$ that can be written in the following way:

$$(\mathcal{D}(g)f)(x) = ((g^{-1})^* f)(x) \ .$$

This also allows to prove unitarity, using $g(\mathbb{R}^3) = \mathbb{R}^3$ and $g^* dV = dV$:

$$\begin{aligned}\langle \mathcal{D}(g)f | \mathcal{D}(g)h \rangle &= \int_{\mathbb{R}^3} f \circ g^{-1} \cdot h \circ g^{-1} \ dV = \int_{\mathbb{R}^3} (f \cdot h) \circ g^{-1} \ dV \\ &= \int_{\mathbb{R}^3} (g^{-1})^* (f \cdot g \ dV) = \int_{g^{-1}(\mathbb{R}^3)} f \cdot h \ dV = \int_{\mathbb{R}^3} f \cdot h \ dV \\ &= \langle f | h \rangle \ .\end{aligned}$$

Since $e^{\varphi X}$ defines a curve in $\mathrm{SO}(3)$ with tangent vector X in $\varphi = 0$, the second claim follows from the previous theorem. □

Corollary D.1.18.
The generators J_i *of Rotations in* $\mathrm{SO}(3)$ *have the following representation on* $L^2(\mathbb{R}^3)$*:*

$$\mathcal{D}_*(J_i) = -\sum_{j,k} \varepsilon_{ijk} x_j \cdot \partial_{x_k}$$

With the quantum mechanical generators $L_\alpha = i\hbar J_\alpha$ *it follows that:*

$$\vec{L} \coloneqq \begin{pmatrix} \mathcal{D}_*(L_1) \\ \mathcal{D}_*(L_2) \\ \mathcal{D}_*(L_3) \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \begin{pmatrix} -i\hbar\partial_{x_1} \\ -i\hbar\partial_{x_2} \\ -i\hbar\partial_{x_3} \end{pmatrix} =: \vec{x} \times \vec{p} \ .$$

Proof D.1.19 (Œ for J_3).
Let $f \in L^2(\mathbb{R}^3)$, then:

$$\mathcal{D}_*(J_3)f = \frac{d}{d\varphi}\bigg|_{\varphi=0} \mathcal{D}(e^{\varphi J_3})f = \frac{d}{d\varphi}\bigg|_{\varphi=0} f \circ e^{-\varphi J_3} = df \circ (-J_3)$$

$$\begin{aligned}\Rightarrow \quad (\mathcal{D}_*(J_3)f)(\vec{x}) &= -df(J_3\vec{x}) = -df(x_1\partial_{x_2} - x_2\partial_{x_1}) \\ &= -x_1 \cdot df(\partial_{x_2}) + x_2 df(\partial_{x_1}) \\ &= x_2 \cdot \frac{\partial}{\partial x_1} f(\vec{x}) - x_1 \cdot \frac{\partial}{\partial x_2} f(\vec{x}) \ .\end{aligned}$$

□

D.2. Some important algebras

Lemma C.1.7 allows to regard tensor products as associative algebras (by isomorphy), defining the **tensor algebra**:

$$\bigotimes V := \bigoplus_{n=0}^{\infty} V^{\otimes n} = K \bigoplus V \bigoplus (V \otimes V) \bigoplus \dots$$

D.2.1. Clifford algebra and Weyl algebra

Definition D.2.1.
Let V be a $\mathbb{K}$ vector space with $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. A **quadratic form** $q\colon V \to \mathbb{K}$ is a map with the following properties:

i) For all $v \in V$ and $\lambda \in \mathbb{K}$ it holds that $q(\lambda v) = \lambda^2 \cdot q(v)$.
ii) The map $Q(v,w) = \frac{1}{2}(q(v+w) - q(v) - q(w))$ is a symmetric bilinear form.

The pair (V, q) is called **quadratic space**.

In the finite-dimensional case, a quadratic form can be expressed by a matrix A, by $q(\vec{x}) = \vec{x}^T A \vec{x}$. Some authors also call the associated symmetric bilinear form Q quadratic form, since

$$Q(v,v) = q(v)$$

holds anyway.

Definition D.2.2.
The **Clifford algebra** $\mathrm{Cl}(V,q)$ is the free associative algebra over $V \oplus \mathbb{K}$ with the restriction

$$v \circ v = q(v) \ .$$

The free associative algebra is the set of all "polynomials" with "variables" from $V \oplus \mathbb{K}$ and coefficients from $\mathbb{K}$. The associative multiplication of variables can be understood as the building of words from an alphabet consisting of $V \oplus \mathbb{K}$:

$$\mathrm{Free}(V) \ni X = \sum_{j=1}^{n} \alpha_j \prod_{\ell} v_{j,\ell} \ , \qquad \alpha_j \in \mathbb{K} \text{ and } v_{j,\ell} \in V \oplus \mathbb{K} \ .$$

Now, the Clifford algebra is the same algebra with the convention, that two successive letters can be exchanged with $q(v)$. The bilinearity of the algebra product allows to pull out this number as coefficient for the "word".

Corollary D.2.3 (Clifford relation for two vectors)**.**
For $v, w \in V$ it holds that:

$$v \circ w + w \circ v = 2Q(v,w) \ .$$

Proof D.2.4.

$$\begin{aligned} q(v+w) &= v \circ v + w \circ w + v \circ w + w \circ v \\ &= q(v) + q(w) + v \circ w + w \circ v \end{aligned}$$

$$\Rightarrow \quad v \circ w + w \circ v = (q(v+w) - q(v) - q(w)) = 2Q(v,w) \ .$$

□

Remark D.2.5.
Following the literature, we will no longer use a special symbol for the Clifford product:

$$e_i \circ e_j \equiv e_i e_j \ .$$

As an algebra, the Clifford algebra carries a vector space structure.

Lemma D.2.6.
Let $\dim(V) = n$ *and* $\{e_1, ..., e_n\}$ *be a basis of* V *that is orthogonal w.r.t.* Q*, then*

$$\{1,\ e_i,\ e_i e_j,\ \ldots,\ e_1 \ldots e_n | i = 1, \ldots, n,\ i < j,\ \ldots\}$$

is a basis of the Clifford algebra $\mathrm{Cl}(V,Q)$*, in the sense of a vector space.*

Proof D.2.7.
Assume the set $V \oplus \mathbb{K}$ to be an alphabet and let $v_1 \ldots v_k$ be a word of length $k \leq n$. Because of the bilinearity, the word can be decomposed into a linear combination of words of the length k built from the sub alphabet $\{1, e_1, \ldots, e_n\}$. Hence this sub alphabet generates all words of length $\leq n$.

Assume now $uv_1 \ldots v_n$ to be a word of length $n+1$. Using $e_i^2 = Q(e_1)$, we see that this word can be decomposed into a linear combination of words with length n. Thus the set

$$\{1,\ e_i,\ e_i e_j,\ \ldots,\ e_1 \ldots e_n\}$$

generates the Clifford algebra.

From lemma D.2.3, the relation between the basis vectors follows immediately:

$$e_i e_j = 2Q(e_i, e_j) - e_j e_i = -e_j e_i \ .$$

Similarly to the exterior algebra it holds that

$$\mathrm{span}\{e_{i_1} \ldots e_{i_k} | i_1 < \ldots < i_k\} = \mathrm{span}\{e_{i_1} \ldots e_{i_k}\} \ ,$$

with the left set being liearly independent. Hence the Clifford algebra is generated by

$$\{1,\ e_i,\ e_i e_j,\ \ldots,\ e_1 \ldots e_n | i = 1, \ldots, n,\ i < j,\ \ldots\}$$

Linear independence follows from the fact, that words of length $k \leq n$ cannot be generated form words shorter than k. Formally one could use an induction here... □

Similarly to the Clifford algebra, that is defined on a vector space with quadratic form, the Weyl algebra is an associative algebra defined on a symplectic vector space. That is a vector space with a non-degenerate alternating bilinear form (2-form).

Definition D.2.8.
Let V be a $\mathbb{K}$ vector space with a non-degenerate 2-form $A : V \times V \to \mathbb{K}$. The **Weyl algebra** $\mathcal{W}(V, A)$ is the free algebra over $V \oplus \mathbb{K}$ with the restriction

$$vw - wv = A(v, w) \qquad \forall\, v, w \in V \ .$$

D.2.2. Exterior algebra and symmetric algebra

Definition D.2.9.
The **symmetric algebra** $S(V)$ is the free commutative algebra over $V \oplus \mathbb{K}$, i.e.:

$$v \vee w = w \vee v \qquad \forall\, v, w \in V \text{ and } \mathbb{K} \ .$$

The **exterior algebra** $\bigwedge(V)$ is the free "alternating" algebra over $V \oplus \mathbb{K}$, i.e.:

$$v \wedge w = -w \wedge v \qquad \forall\, v, w \in V \text{ and } \mathbb{K} \ .$$

All of the algebras we have introduced so far can also be constructed by the tensor algebra modulus some ring-ideal, which would also prove existence. In the case of the symmetric and exterior algebra it can be seen, that the tensor algebra can be decomposed into a direct sum $\bigotimes(V) = \bigwedge(V) \oplus S(V)$.

The algebras $\bigwedge(V)$ and $S(V)$ are graded. This means, that there are sub-algebras $\bigwedge^k(V) \subset \bigwedge(V)$ and $S^k(V) \subset S(V)$, such that

$$\bigwedge(V) = \bigoplus_{k=0}^{\infty} \bigwedge^k(V) \qquad \text{and} \qquad S(V) = \bigoplus_{k=0}^{\infty} S^k(V) \ .$$

The grade k is the length of the polynomial $v_1 \wedge \ldots \wedge v_k$ (resp. $v_1 \vee \ldots \vee v_k$). Since $v \wedge v = -v \wedge v = 0$, polynomials of $\bigwedge(V)$ can only be of length $n = \dim(V)$, such that

$$\bigwedge(V) = \bigoplus_{k=0}^{n} \bigwedge^k(V) \ .$$

Lemma D.2.10.
Let $\dim(V) = n$ *and* $\{e_1, ..., e_n\}$ *be a basis of* V*, then*

$$\{e_{i_1} \wedge \ldots \wedge e_{i_k} \mid i_1 < i_2 \ldots < i_k\}$$

is a basis of the exterior algebra $\bigwedge^k(V)$*, in the sense of a vector space.*

Proof D.2.11.
Consider a polynomial $v_1 \wedge \ldots \wedge v_k \neq 0$ of length $k \leq n$. Note that $v_i = v_j$ for

i, j leads to $v_1 \wedge \ldots \wedge v_k = 0$ because of $v \wedge v = -v \wedge v$. Because of alternating multilinearity, it follows that polynomials of length k can be written as linear combination of $e_{i_1} \wedge \ldots e_{i_k}$ for $i_j \neq i_k$ and $i_j \in \{1, \ldots, n\}$. Hence the set

$$\{e_{i_1} \wedge \ldots \wedge e_{i_k} \mid i_1 \neq i_2 \neq \ldots \neq i_k\}$$

generates $\bigwedge^k(V)$. Because of the alternation, we can exchange the order, possibly getting a minus sign, to order the e_i, such that $i_1 < i_2 \ldots < i_k$. Thus

$$\{e_{i_1} \wedge \ldots \wedge e_{i_k} \mid i_1 < i_2 \ldots < i_k\}$$

is a basis of $\bigwedge^k(V)$. □

Lemma D.2.12.
Let $\dim(V) = n$ *and* $\{e_1, ..., e_n\}$ *be a basis of* V*, then*

$$\{1,\ e_i,\ e_i e_j,\ \ldots,\ e_1 \ldots e_n | i = 1, \ldots, n,\ i \leq j,\ \ldots\}$$

is a basis of the exterior algebra $S(V)$*, in the sense of a vector space.*

Proof D.2.13.
The proof works in the same way as the one for the exterior algebra. The only addition is, that $v \vee v$ has to be admitted, such that $i_1 \leq i_2 \leq \ldots \leq i_k$ has to be used in the last step. □

For the physics, one is in general also interested in $\bigwedge(V \oplus W)$ and $S(V \oplus W)$.

Lemma D.2.14.
For finite dimensional vector spaces, it holds that

$$\bigwedge^p(V \oplus W) \cong \bigoplus_{q=0}^{p} \bigwedge^q(V) \otimes \bigwedge^{p-q}(W)\ ,$$

where the isomorphy is on the vector space level.

Proof D.2.15.
Let $\{e_i\}_{i=1,\ldots,n}$ be a basis of V and $f_{j\,j=1,\ldots,m}$ be a basis of W. Then

$$\{(e_1, 0), \ldots, (e_n, 0), (0, f_1), \ldots, (0, f_m)\}$$

is a basis of $V \oplus W$. It is convenient to write $(v, w) = v + w$ here, identifying $(e_i, 0) = e_i$ and $(0, f_j) = f_j$. Furthermore, we change the notation for a moment, and set $e_{n+j} = f_j$. From lemma D.2.10, we know that

$$\{e_{i_1} \wedge \ldots \wedge e_{i_p} \mid i_k \in \{0, 1, \ldots m + n\}\ ,\ i_1 < \ldots < i_p\}$$

is a basis of $\bigwedge^p(V \oplus W)$. So we see, that the basis of $\bigwedge^p(V \oplus W)$ has the form

$$\{(e_{i_1} \wedge \ldots \wedge e_{i_q}) \wedge (f_{j_1} \wedge \ldots \wedge f_{j_{p-q}}) \mid q \in \{0, \ldots, p\} \ , \ i_1 < \ldots < i_q \ , \ j_1 < \ldots < j_{p_{p-q}}\} \ .$$

On the other hand, by theorem C.1.5,

$$\{(e_{i_1} \wedge \ldots \wedge e_{i_q}) \otimes (f_{j_1} \wedge \ldots \wedge f_{j_{p-q}}) \mid i_1 < \ldots < i_q \ , \ j_1 < \ldots < j_{p_{p-q}}\}$$

is a basis of $\bigwedge^q(V) \otimes \bigwedge^{p-q}(W)$. In the same way, we constructed the basis for $V \oplus W$, we can construct the basis for $\bigoplus_{q=0}^p \bigwedge^q(V) \otimes \bigwedge^{p-q}(W)$. Up to reordering, we find, that

$$\{(e_{i_1} \wedge \ldots \wedge e_{i_q}) \otimes (f_{j_1} \wedge \ldots \wedge f_{j_{p-q}}) \mid q \in \{0, \ldots, p\} \ , \ i_1 < \ldots < i_q \ , \ j_1 < \ldots < j_{p_{p-q}}\}$$

is the basis. So we can construct the isomorphism $\phi\colon \bigwedge^p(V \oplus W) \xrightarrow{\cong} \bigoplus_{q=0}^p \bigwedge^q(V) \otimes \bigwedge^{p-q}(W)$ by its action on the basis elements:

$$\phi\colon (e_{i_1} \wedge \ldots \wedge e_{i_q}) \wedge (f_{j_1} \wedge \ldots \wedge f_{j_{p-q}}) \longmapsto (e_{i_1} \wedge \ldots \wedge e_{i_q}) \otimes (f_{j_1} \wedge \ldots \wedge f_{j_{p-q}}) \ .$$

Extending the map linearly makes ϕ a linear isomorphism. □

Lemma D.2.16.
For finite dimensional vector spaces, it holds that

$$S^p(V \oplus W) \cong \bigoplus_{q=0}^{p} S^q(V) \otimes S^{p-q}(W) \ ,$$

where the isomorphy is on the vector space level.

Proof D.2.17.
The proof is the same as for the exterior algebra, except for $\leq$ instead of $<$. □

Corollary D.2.18.
The algebras $\bigwedge(V \oplus W)$ and $S(V \oplus W)$ are bigraded with

$$\bigwedge(V \oplus W) = \bigoplus_{p,q} \bigwedge^{p,q}(V \oplus W) \ , \qquad \textit{where} \quad \bigwedge^{p,q}(V \oplus W) \equiv \bigwedge^p(V) \otimes \bigwedge^q(W) \ ,$$

$$S(V \oplus W) = \bigoplus_{p,q} S^{p,q}(V \oplus W) \ , \qquad \textit{where} \quad S^{p,q}(V \oplus W) \equiv S^p(V) \otimes S^q(W) \ .$$

Proof D.2.19.
This follows from the previous lemmas and the grading of $\bigwedge$ and S. □

D.3. Spinor representation

Let $V \simeq \mathbb{R}^{2n}$ be an even dimensional real vector space with Euclidean scalar product Q. Since the dimension is even, the basis can be split up in two parts of equal length. Let $\{e_1, \dots, e_n, f_1, \dots, f_n\}$ be an orthonormal basis. To allow for complex coefficients, we use the complexification $V^{\mathbb{C}}$ (see definition C.1.9). To represent the Clifford algebra, we intend to use something similar to a polarization of a symplectic space. That is, we want to define subspaces P and P^*, such that $V^{\mathbb{C}} = P \oplus P^*$. One possibility, using the chosen basis, is the construction

$$P = \operatorname{span}_{\mathbb{C}}\{c_1, \dots, c_n\} \quad P^* = \operatorname{span}_{\mathbb{C}}\{c_1^*, \dots, c_n^*\}$$

$$\text{with} \qquad c_j = \tfrac{1}{2}(e_j - if_j)\ , \quad c_j^* = \tfrac{1}{2}(e_j + if_j)\ .$$

The Euclidean scalar product allows to regard P^* as dual space, since $Q(c_k^*, c_\ell) = \delta_{k\ell}$. Also, it holds that $Q(c_k, c_\ell) = 0 = Q(c_k^*, c_\ell^*)$. Passing to the Clifford algebra $\mathrm{Cl}(V^{\mathbb{C}}, Q)$, the relations for the basis elements of P and $P*$ are

$$c_k c_\ell + c_\ell c_k = 0\ , \qquad c_k^* c_\ell^* + c_\ell^* c_k^* = 0\ , \qquad c_k^* c_\ell + c_\ell c_k^* = \delta_{k\ell}\ .$$

An important observation is, that $\mathbb{C}$ and P are both subsets of $\mathrm{Cl}(V^{\mathbb{C}}, Q)$ and $\bigwedge(P)$.

Definition D.3.1.
Let $\varphi \in P^*$, the linear operator $\varphi\lrcorner\colon \bigwedge^k(P) \to \bigwedge^{k-1}(P)$ is defined by

i) $\varphi\lrcorner a = 0$ for $a \in \mathbb{C}$.

ii) $\varphi\lrcorner v = \varphi(v) \coloneqq 2Q(\varphi, v)$ for $v \in P$

iii) $\varphi\lrcorner$ satisfies the anti Leibniz rule: $\varphi\lrcorner(\xi \wedge \eta) = (\varphi\lrcorner\xi) \wedge \eta + (-1)^{\deg(\xi)} \xi \wedge (\varphi\lrcorner\eta)$.

The operator

$$v\wedge\colon \bigwedge^k(P) \to \bigwedge^{k+1}(P)\ ,$$

for $v \in P$, is defined as the notation implies. Since every element in $v \in V^{\mathbb{C}} \oplus \mathbb{C}$ can be uniquely decomposed into $v = v_P + v_{P^*} + a$ with $v_P \in P$, $v_{P^*} \in P^*$ and $a \in \mathbb{C}$, an operator $v\circ\colon \bigwedge(P) \to \bigwedge(P)$ can be defined by

$$v \circ \xi = v_P \wedge \xi + v_{P^*}\lrcorner\xi + a \cdot \xi \qquad \forall\ \xi \in \bigwedge(P)\ .$$

Demanding that

$$(vw) \circ \xi = v \circ (w \circ \xi) \quad \text{and} \quad (a \cdot v + b \cdot w) \circ \xi = a \cdot v \circ \xi + b \cdot w \circ \xi\ ,$$

for all $v, w \in V^{\mathbb{C}} \oplus \mathbb{C}$, $\xi \in \bigwedge(P)$ and $a, b \in \mathbb{C}$, allows to extend $v\circ$ for any $v \in \mathrm{Cl}(V^{\mathbb{C}}, Q)$. Also by construction $v\circ \in \mathrm{End}(\bigwedge(P))$.

Lemma D.3.2.
The map $\mathcal{D}_S\colon v \in \mathrm{Cl}(V^{\mathbb{C}}, Q) \to \mathrm{End}(\bigwedge(P))$ defined by $v \mapsto v\circ$ is an algebra homomorphism.

Proof D.3.3.
In fact, $v\circ$ has been constructed such that $\mathcal{D}_S$ is linear. The homomorphy follows equally simply from the construction of $v\circ$:

$$\mathcal{D}_S(vw)(\xi) = (vw) \circ \xi = v \circ (w \circ \xi) = \mathcal{D}_S(v)(\mathcal{D}_S(w)\xi) = (\mathcal{D}_S(v)\mathcal{D}_S(w))\xi \ .$$

□

With the algebra homomorphism $\mathcal{D}_S$ at hand, we are ready to define the spinor representation.

Definition D.3.4.
Let $V^{\mathbb{C}} = P \oplus P^*$ be a polarization of the complexified vector space $V^{\mathbb{C}}$. The **spinor representation** of $\mathrm{Cl}(V^{\mathbb{C}}, Q)$ on $\bigwedge(P)$ is $\mathcal{D}_S$.

In the same way, the spinor representation can be constructed on $\bigwedge(P^*)$

D.4. Bosonic and fermionic Fock space representations

Essentially, the following representations are extensions of the spinor representation for Clifford and Weyl algebras on Hilbert spaces. The following lemma allows these extensions:

Lemma D.4.1.
Let $\mathcal{H}$ be a Hilbert space and $\mathcal{H}^*$ its dual. Then $\mathcal{H} \oplus \mathcal{H}^*$ is a vector space with

i) A symmetric bilinar form $Q(v + \varphi, v' + \varphi') = \varphi(v') + \varphi'(v)$,

ii) A symplectic form $A(v + \varphi, v' + \varphi') = \varphi(v') - \varphi'(v)$,

for $v, v' \in \mathcal{H}$ and $\varphi, \varphi' \in \mathcal{H}^*$.

Proof D.4.2.
The vector space property follows from the construction of $\mathcal{H} \oplus \mathcal{H}^*$. Also, symmetry and bilinearity of Q can be seen easily, as well as bilinearity and alternation of A. It remains to show the non-degeneracy of A.

Assume $A(v + \varphi, v' + \varphi') = 0$ for all $v + \varphi \in \mathcal{H} \oplus \mathcal{H}^*$. Then:

$$\varphi(v') = \varphi'(v) \quad \forall\ v + \varphi \in \mathcal{H} \oplus \mathcal{H} * \ .$$

Assume $\varphi \neq 0$ and $v = 0$, then $\varphi(v') = 0$, which can only be true, if $v' = 0$. Conversely for $\varphi = 0$ and $v \neq 0$, then $\varphi'(v) = 0$ is needed, which is only possible for all non zero v, if $\varphi' = 0$. Hence $v' + \varphi' = 0$. □

D.4.1. Fermionic Fock space representation

Let $\mathcal{H}$ be a Hilbert space with dual space $\mathcal{H}^*$. The **fermionic Fock space representation** $\mathcal{D}_F$ is the representation of $\mathrm{Cl}(\mathcal{H} \oplus \mathcal{H}^*, Q)$ on $\bigwedge(\mathcal{H})$, defined by $\mathcal{D}_F(v) = v\circ$ for all $v \in \mathrm{Cl}(\mathcal{H} \oplus \mathcal{H}^*, Q)$. The action $v\circ$ is defined as in the spinor representation and the symmetric bilinear form is the one from lemma D.4.1.

Definition D.4.3.
The **anticommutator** $\{\cdot,\cdot\}$ of two operators $A, B \in \mathrm{End}(\mathcal{H})$ is defined by

$$\{A, B\} = AB + BA \ .$$

Theorem D.4.4.
Let $v, w \in \mathcal{H}$ and $\varphi, \eta \in \mathcal{H}^*$, then the representations of these elements satisfy the ***canonical anticommutator relations****:*

$$\{\mathcal{D}_F(v), \mathcal{D}_F(w)\} = 0 \ , \qquad \{\mathcal{D}_F(\varphi), \mathcal{D}_F(\eta)\} = 0 \ ,$$

$$\{\mathcal{D}_F(v), \mathcal{D}_F(\varphi)\} = \varphi(v) \ .$$

Proof D.4.5.

i) The first anticommutator rule is a direct consequence of $v \wedge w = -w \wedge v$, since $\mathcal{D}_F(v)\mathcal{D}_F(w)\xi = v \wedge w \wedge \xi$ for all $\xi \in \bigwedge(\mathcal{H})$.

ii) Because of linearity, we can restrict to $\xi = x_1 \wedge \ldots \wedge x_m$. Using $\mathcal{D}_F(\varphi)v = \varphi \lrcorner v = \varphi(v)$ and the anti Leibniz rule, we get

$$\mathcal{D}_F(\varphi)\xi = \sum_{j=1}^{m} (-1)^{j+1} \varphi(x_j) \cdot x_1 \wedge \ldots \wedge \widehat{x}_j \wedge \ldots \wedge x_m \ ,$$

where $\widehat{x}_j$ is the omission of x_j. This results in

$$\begin{aligned}
\mathcal{D}_F(\eta)\mathcal{D}_F(\varphi)\xi &= \sum_{j=1}^{m} \sum_{k=1, k\neq j}^{m} (-1)^{j+k} \varphi(x_j)\eta(x_k) \\
&\qquad \cdot x_1 \wedge \ldots \wedge \widehat{x}_j \wedge \ldots \wedge \widehat{x}_k \wedge \ldots \wedge x_m \\
&= \sum_{j=1}^{m} \sum_{k=1, k\neq j}^{m} (-1)^{\Delta j + \Delta k} (-1)^{j+k} \varphi(x_k)\eta(x_j) \\
&\qquad \cdot x_1 \wedge \ldots \wedge \widehat{x}_k \wedge \ldots \wedge \widehat{x}_j \wedge \ldots \wedge x_m \\
&= (-1)^{\Delta j + \Delta k} \mathcal{D}_F(\varphi)\mathcal{D}_F(\eta)\xi \ .
\end{aligned}$$

In the second line, x_j and x_k where exchanged. This is possible, for the summation is over all possible combinations. However, exchanging x_j and x_k results in an additional sign $(-1)^{\Delta j + \Delta k}$. Let Δk be the number of pair exchanges, to bring x_k in front of x_j. The number of pair exchanges Δj to

bring x_j to the former position of x_k is then one less that Δk, i.e. $\Delta j = \Delta k - 1$. It follows that

$$(-1)^{\Delta j + \Delta k} = (-1)^{2\Delta k - 1} = -1 \ .$$

Hence: $\mathcal{D}_F(\eta)\mathcal{D}_F(\varphi) = -\mathcal{D}_F(\varphi)\mathcal{D}_F(\eta)$.

iii) Again, let $\xi = x_1 \wedge \ldots \wedge x_m$, then

$$\begin{aligned}
&\mathcal{D}_F(v)\mathcal{D}_F(\varphi)\xi + \mathcal{D}_F(\varphi)\mathcal{D}_F(v)\xi \\
&\quad = \sum_{j=1}^{m} (-1)^{j+1} \varphi(x_j) \cdot v \wedge x_1 \wedge \ldots \wedge \widehat{x}_j \wedge \ldots \wedge x_m \\
&\quad + \sum_{j=1}^{m} (-1)^{j+2} \varphi(x_j) \cdot v \wedge x_1 \wedge \ldots \wedge \widehat{x}_j \wedge \ldots \wedge x_m \\
&\quad + \varphi(v) \cdot x_1 \wedge \ldots \wedge x_m \\
&\quad = \varphi(v) \cdot \xi \ .
\end{aligned}$$

□

D.4.2. Bosonic Fock space representation

Let $\mathcal{H}$ be a Hilbert space with dual space $\mathcal{H}^*$ and A the symplectic form on $\mathcal{H} \oplus \mathcal{H}^*$ from lemma D.4.1. The **bosonic Fock space representation** $\mathcal{D}_B$ is a representation of $\mathcal{W}(\mathcal{H} \oplus \mathcal{H}^*, A)$ on $S(\mathcal{H})$ similar to the fermionic Fock space representation. As before $\mathcal{D}_B$ is defined by $\mathcal{D}_B(v) = v\circ$ for all $v \in \mathcal{W}(\mathcal{H} \oplus \mathcal{H}^*, A)$. However, the action $v\circ$ on $S(\mathcal{H})$ needs some minor changes:

$$w \circ \xi \coloneqq w \vee \xi \qquad \forall\ \xi \in S(\mathcal{H})\ ,\ w \in \mathcal{H}$$

and $\varphi\lrcorner$ satisfies the Leibniz rule: $\varphi\lrcorner(\xi \vee \eta) = (\varphi\lrcorner\xi) \vee \eta + \xi \vee (\varphi\lrcorner\eta)$.

for all $\varphi \in \mathcal{H}^*$ and all $\xi, \eta \in S(\mathcal{H})$.

Theorem D.4.6.
Let $v, w \in \mathcal{H}$and $\varphi, \eta \in \mathcal{H}^*$, then the representations of these elements satisfy the ***canonical commutator relations****:*

$$[\mathcal{D}_B(v), \mathcal{D}_B(w)] = 0\ , \qquad [\mathcal{D}_B(\varphi), \mathcal{D}_B(\eta)] = 0\ ,$$

$$[\mathcal{D}_B(v), \mathcal{D}_B(\varphi)] = \varphi(v)\ .$$

Proof D.4.7.
The proof is very similar to proof D.4.5, yet with less sign changes. □

D.5. Spin Group

For the whole section we assume the quadratic form Q to be non-degenerate, if not specified differently.

D.5.1. Spin and pinor group

Definition D.5.1.
Let (V, Q) be a quadratic space. The **pinor group** $\mathrm{Pin}(V, Q)$ is defined by

$$\mathrm{Pin}(V,Q) = \left\{ X \in \mathrm{Cl}(V,Q) \;\middle|\; X = v_1 \dots v_k,\ k \geq 0,\ v_\ell \in V,\ \text{with } Q(v_\ell, v_\ell)^2 = 1 \right\} .$$

From the definition, it is not hard to see, that $\mathrm{Pin}(V, Q) \subset \mathrm{Cl}(V, Q)$ and that it is closed under the Clifford multiplication. The pinor group can be understood as a restriction of the Clifford algebra.

From the universal property of Clifford algebras, which is itself sufficient to completely characterize Clifford algebras abstractly, it can be shown that there is a canonical isomorphism $\alpha \colon \mathrm{Cl}(V, Q) \to \mathrm{Cl}(V, Q)$. However, since we did not use the universal property, we will characterize α using a Q-orthonormal basis $\{e_j\}$ and extend it linearly:

$$\alpha(e_{j_1} \dots e_{j_k}) = (-1)^k e_{j_1} \dots e_{j_k} \quad \text{for } 1 \leq j_1 < \dots < j_k \leq \dim(V) .$$

ChapterStart

Definition D.5.2.
The even/odd Clifford sub algebras $\mathrm{Cl}^+(V, Q)$ and $\mathrm{Cl}^-(V, Q)$ are defined by

$$\mathrm{Cl}^+(V,Q) \coloneqq \{X \in \mathrm{Cl}(V,Q) \mid \alpha(X) = X\}$$

$$\text{and} \quad \mathrm{Cl}^-(V,Q) \coloneqq \{X \in \mathrm{Cl}(V,Q) \mid \alpha(X) = -X\} .$$

Example D.5.3.
Consider $\mathrm{Cl}_2(V, Q)$. Every element $X \in \mathrm{Cl}_2(V, Q)$ can be written as $X = \sum_{j<k} X_{jk} e_j e_k$ using the basis from lemma D.2.6. Since $e_j e_k = -e_k e_j$ for $j \neq k$ on can also write

$$X = \sum_{j,k} X_{jk} e_j e_k \qquad \text{where } X_{jk} = -X_{kj} .$$

It is clear, that every $X = \sum_j (u_j v_j - v_j u_j) = \sum_j [u_j, v_j]$ with $u_j, v_j \in V$ is an element of $\mathrm{Cl}_2(V, Q)$, since

$$\begin{aligned} uv - vu &= \sum_{i,j} u_i v_j (e_i e_j - e_j e_i) = 2 \sum_{i,j \neq i} u_i v_j e_i e_j \\ &= 2 \sum_{i<j} (u_i v_j - u_j v_i) e_i e_j \in \mathrm{Cl}_2(V,Q). \end{aligned}$$

But also the opposite is true:

$$\sum_{j<k} X_{jk} e_j e_k = \frac{1}{2} \sum_{j,k\neq j} X_{jk}(e_j e_k - e_k e_j) = \sum_j (e_j v_j - v_j e_j) = \sum_j [e_j, v_j] \ ,$$

with $v_j = \frac{1}{2}\sum_{k\neq j} X_{jk} e_k$. However this means

$$\mathrm{Cl}_2(V,Q) = \{X \in \mathrm{Cl}(V,Q) \mid X = \sum_j [u_j, v_j] \text{ with } u_j, v_j \in V\} \ .$$

Lemma D.5.4.
The sub algebra $Cl_2(V,Q)$ is a Lie algebra w.r.t. the commutator.

Proof D.5.5.
The commutator satisfies all of the conditions to be a Lie bracket. We need only show that $\mathrm{Cl}_2(V,Q)$ is closed under the commutator. In fact, because of the linearity of the commutator, it is sufficient to show $[e_i e_j, e_k e_\ell] \in \mathrm{Cl}_2(V,Q)$ for $i<j$ and $k<\ell$:

$$\begin{aligned}
[e_i e_j, e_k e_\ell] &= e_i e_j e_k e_\ell - e_k e_\ell e_i e_j = -e_i e_k e_j e_\ell + 2\delta_{jk} e_i e_\ell - e_k e_\ell e_i e_j \\
&= e_k e_i e_j e_\ell - 2\delta_{ik} e_j e_\ell + 2\delta_{jk} e_i e_\ell - e_k e_\ell e_i e_j \\
&= \ldots = e_k e_\ell e_i e_j - e_k e_\ell e_i e_j - 2\delta_{i\ell} e_k e_j + 2\delta_{j\ell} e_k e_i - 2\delta_{ik} e_j e_\ell + 2\delta_{jk} e_i e_\ell \\
&= -2\delta_{i\ell} e_k e_j + 2\delta_{j\ell} e_k e_i - 2\delta_{ik} e_j e_\ell + 2\delta_{jk} e_i e_\ell \in \mathrm{Cl}_2(V,Q) \ .
\end{aligned}$$

□

Definition D.5.6.
The **spin group**, $\mathrm{Spin}(V,Q)$ is defined as

$$\mathrm{Spin}(V,Q) = \mathrm{Pin}(V,Q) \cap \mathrm{Cl}^+(V,Q) \ .$$

Definition D.5.7.
The Lie algebra $\mathrm{Cl}_2(V,Q)$ is called **spin algebra**, $\mathfrak{spin}(V,Q)$.

Assigning both the algebra $\mathrm{Cl}_2(V,Q)$ and the group $\mathrm{Pin}(V,Q) \cap \mathrm{Cl}^+(V,Q)$ with *spin*, is no coincidence:

Theorem D.5.8 (see [Fri03, Satz 4.2])**.**
The Lie algebra of Spin(V,Q), for a real quadratic space (V,Q) is $\mathfrak{spin}(V,Q)$.

The proof of this theorem is not trivial and requires more Lie theory and topology. On the other hand we could understand $\mathrm{Spin}(V,Q)$ as the Lie algebra corresponding to $\mathrm{Cl}_2(V,Q)$, that results from exponentiation in the Clifford algebra. Indeed, **Lie's third theorem** states, that any finite-dimensional real Lie algebra is the Lie algebra of a Lie

group. The mentioning of exponentiation inside the Clifford algebra is important, since different Lie groups may have the same Lie algebra.

D.5.2. Relation to the special orthogonal group

The spin group and the special orthogonal group are one example of the aforementioned fact, that different Lie groups can have the same Lie algebra. But first we have to show, that these Lie groups have the same Lie algebra:

Lemma D.5.9.
The map $\tau\colon \mathfrak{spin}(V,Q) \to \mathfrak{so}(V,Q)$ *defined by* $\tau(X) = [X,\cdot]$ *is a Lie algebra isomorphism.*

Proof D.5.10.
Before we can show, that the map is an isomorphism, we need to show that it is a well defined map in the first place. Thus we need to show, that for all $v \in V$ the element $[X,v] \in \mathrm{Cl}(V,Q)$ is indeed an element of V:

$$\begin{aligned}
[X,v] &= \left[\sum_j (x_j y_j - y_j x_j), v\right] = \sum_j ([x_j y_j, v] - [y_j x_j, v]) \\
&= \sum_j (2[x_j y_j, v] - 2Q(x_j, y_j)[1, v]) \\
&= 2\sum_j [x_j y_j, v] = 2\sum_j (x_j y_j v - v x_j y_j) \\
&= 2\sum_j (x_j v y_j + 2Q(y_j, v) - x_j v y_j - 2Q(x_j, v) y_j) \\
&= 4\sum_j (Q(v, y_j) x_j - Q(v, x_j) y_j) \in V \ .
\end{aligned}$$

To be an element of $\mathfrak{so}(V,Q)$, the condition $Q(\tau(X)v, w) + Q(v, \tau(X)w) = 0$ has to be satisfied:

$$\begin{aligned}
&Q(\tau(X)v, w) + Q(v, \tau(X)w) \\
&\qquad = Q([X,v], w) + Q([X,w], v) \\
&\qquad = 4\sum_j (Q(v, y_j)Q(w, x_j) - Q(v, x_j)Q(w, y_j) \\
&\qquad\qquad + Q(w, y_j)Q(v, x_j) - Q(w, x_j)Q(v, y_j)) \\
&\qquad = 0 \ .
\end{aligned}$$

The map τ is a Lie algebra homomorphism:

$$\begin{aligned}
\tau([X,Y])v &= [[X,Y], v] = -[[v,X], Y] - [[Y,v], X] \\
&= [Y, -\tau(X)v] + [X, \tau(Y)v] = \tau(X)\tau(Y)v - \tau(Y)\tau(X)v \\
&= [\tau(X), \tau(Y)]v \ .
\end{aligned}$$

Since the map τ is linear between finite-dimensional vector spaces, bijectivity is equivalent to injectivity (rank nullity theorem), which is equivalent to a trivial kernel. Let $\{e_i\}$ be a Q-orthonormal basis, then X can be written as $X = \sum_{i<j} X_{ij} e_i e_j$. Before we show, that $\ker(\tau) = 0$, we calculate the following

$$[e_i e_j, e_k] = Q(e_k, e_j)e_i - Q(e_k, e_i)e_j = \delta_{kj} e_i - \delta_{ki} e_j \ .$$

Assume $\tau(X) = 0$, then $\tau(X)v = 0$ for all $v = \sum_k v_k e_k \in V$. This is equivalent to

$$0 = \sum_{i<j} \sum_k X_{ij} v_k [e_i e_j, e_k] = \sum_{i<j} \sum_k X_{ij} v_k (\delta_{kj} e_i - \delta_{ki} e_j)$$

$$\Rightarrow \qquad \sum_{i<j} X_{ij} v_j e_i = \sum_{i<j} X_{ij} v_i e_j \ .$$

The former is only possible if $X_{ij} \equiv 0$, due to the linear independence of $\{e_i\}$. To see that, choose $i = k$ on the left side and $j = k$ on the right side, then it must hold that

$$\sum_{k<j} X_{kj} v_j = \sum_{i<k} X_{ik} v_i = \sum_{j<k} X_{jk} v_j \ .$$

Since v is arbitrary, and the left side only considers components v_j for $j > k$ while the right side only considers components with $j < k$, it follows that $X_{jk} \equiv 0$, which is equivalent to $X = 0$. □

Since τ is an isomorphism, there has to be an inverse. To prove the construction for the inverse map, we need the following corollary:

Corollary D.5.11.
For a Q-orthogonal basis $\{e_\alpha\}$ it holds that

$$[[e_\alpha, e_\beta], e_\nu] = 4(\delta_{\beta\nu} e_\alpha - \delta_{\nu\alpha} e_\beta)$$

in the Clifford algebra.

Proof D.5.12.

$$\begin{aligned}
[[e_\alpha, e_\beta], e_\nu] &= e_\alpha e_\beta e_\nu - e_\beta e_\alpha e_\nu - e_\nu e_\alpha e_\beta + e_\nu e_\beta e_\alpha \\
&= e_\alpha e_\beta e_\nu - 2\delta_{\alpha\beta} e_\nu + e_\alpha e_\beta e_\nu - e_\nu e_\alpha e_\beta + 2\delta_{\alpha\beta} e_\nu - e_\nu e_\alpha e_\beta \\
&= 2(e_\alpha e_\beta e_\nu - e_\nu e_\alpha e_\beta) \\
&= 2(2\delta_{\beta\nu} e_\alpha - e_\alpha e_\nu e_\beta - 2\delta_{\nu\alpha} + e_\alpha e_\nu e_\beta) \\
&= 4(\delta_{\beta\nu} e_\alpha - \delta_{\nu\alpha} e_\beta) \ .
\end{aligned}$$

□

Lemma D.5.13.

Let $X = X^\alpha{}_\nu e_\alpha \otimes e^\nu \in \mathfrak{so}(V,Q)$, then it holds that

$$\tau^{-1}(X) = \frac{1}{8} X^{\alpha\beta}[e_\alpha, e_\beta] = \frac{1}{8} X_{\alpha\beta}[e^\alpha, e^\beta] \ ,$$

where $X^{\alpha\beta} = X^\alpha{}_\nu Q^{\nu\beta}$.

Proof D.5.14.
The second equality is just the usual raising and lowering of indices, using the linearity of commutator and Clifford product. To probe the first equality, it is easier to prove:

$$X = \tau(\tfrac{1}{8} X^{\alpha\beta}[e_\alpha, e_\beta])$$
$$\Leftrightarrow \qquad Xv = X^\alpha{}_\nu v^\nu e_\alpha = [\tfrac{1}{8} X^{\alpha\beta}[e_\alpha, e_\beta], v] \ .$$

Before we start the calculation, we observe that $X \in \mathfrak{so}(V,Q)$ means that X is skew:

$$Q(Xv, w) + Q(v, Xw) = 0$$
$$\Leftrightarrow \qquad 0 = Q_{\alpha\nu} X^\alpha{}_\beta v^\beta w^\nu + Q_{\alpha\nu} v^\beta X^\beta{}_\nu w^\nu = X_{\nu\beta} v^\beta w^\nu + X_{\beta\nu} v^\beta w^\nu$$
$$\Leftrightarrow \qquad X^\beta{}_\nu = -X^\beta{}_\nu \qquad \Leftrightarrow \qquad X_{\nu\beta} = -X_{\beta\nu} \qquad \Leftrightarrow \qquad X^{\nu\beta} = -X^{\beta\nu} \ .$$

With this relation and corollary D.5.11 we calculate:

$$\begin{aligned}
[\tfrac{1}{8} X^{\alpha\beta}[e_\alpha, e_\beta], v] &= \tfrac{1}{8} X^{\alpha\beta} v^\nu [[e_\alpha, e_\beta], e_\nu] = \tfrac{1}{2} X^{\alpha\beta} v^\nu (\delta_{\beta\nu} e_\alpha - \delta_{\nu\alpha} e_\beta) \\
&= \tfrac{1}{2} (X^\alpha{}_\nu v^\nu e_\alpha - X_\nu{}^\beta v^\nu e_\beta) = \tfrac{1}{2} (X^\alpha{}_\nu v^\nu e_\alpha - X_\nu{}^\alpha v^\nu e_\alpha) \\
&= X^\alpha{}_\nu v^\nu e_\alpha = Xv \ .
\end{aligned}$$

□

In the following we will sketch the relation between $\mathrm{Spin}(V,Q)$ and $\mathrm{SO}(V,Q)$. A rigorous relation needs more topology and Lie theory, as was the case for the proof of theorem D.5.8.

Let $X \in \mathfrak{spin}(V,Q)$ and $v \in V$. A generalization of lemma B.1.4 for the Lie theory allows to write:

$$e^X v e^{-X} = \sum_{n=0}^{\infty} \frac{1}{n!} [X, v]_n = \sum_{n=0}^{\infty} \frac{1}{n!} \tau(X)^n v = e^{\tau(X)} v \ .$$

Since $\tau(X) \in \mathfrak{so}(V,Q)$ and the exponential function is the exponential map for both $\mathrm{SO}(V,Q)$ and $\mathrm{Spin}(V,Q)$, e^X and e^{-X} are elements of $\mathrm{Spin}(V,Q)$ and $e^{\tau(X)}$ is an element of $\mathrm{SO}(V,Q)$. It can be shown, that for both Lie groups the exponential map is surjective. We can define a homomorphism

$$\rho \colon \mathrm{Spin}(V,Q) \longrightarrow \mathrm{SO}(V,Q) \ , \qquad \rho(g) v = g v g^{-1} \ ,$$

such that the following diagram commutes:

$$\begin{array}{ccc} \mathfrak{spin}(V,Q) & \xrightarrow{\exp} & \mathrm{Spin}(V,Q) \\ \tau\downarrow & & \downarrow\rho \\ \mathfrak{so}(V,Q) & \xrightarrow{\exp} & \mathrm{SO}(V,Q) \end{array}$$

Proof D.5.15.
First we show, that the diagram commutes, i.e. $\rho \circ \exp = \exp \circ \tau$:

$$((\rho \circ \exp)X)(v) = \rho(e^X)v = e^X v (e^X)^{-1} = e^X v e^{-X} = e^{\tau(X)}v = ((\exp \circ \tau)X)v \ .$$

The homomorphy follows equally direct:

$$\rho(gh)v = (gh)v(gh)^{-1} = ghvh^{-1}g^{-1} = \rho(g)(hvh^{-1}) = (\rho(g) \circ \rho(h))(v) \ .$$

□

E

The Hilbert space $L^2(\mathbb{R}^n)$

In quantum mechanics the Hilbert space $L^2(\mathbb{R}^n)$ may be the most known Hilbert space, as the representation of wave mechanics uses this hilbert space. In this chapter we want to introduce the physical conventions regarding this space, as can be found in [RW08] for example. On the other hand we want to give fair waring about the mathematical subtleties of these conventions as is done in [Zir13]. The conclusion of this chapter will be the spherical harmonics. We will follow [Zir16] very closely, as it provides an introduction that gets rid of the usually annoying ansatz for solving the partial differential equations.

The set of square integrable functions $\mathcal{L}^2(\mathbb{R}^n)$ is defined by

$$\mathcal{L}^2(\mathbb{R}^n) \coloneqq \{f : \mathbb{R}^n \to \mathbb{C} | f \text{ is integrable and } \int_{\mathbb{R}^n} |f(x)|^2 \, dx^n \quad < \infty\} \, .$$

Using the equivalence relation $f \sim g$, whenever f and g are equal up to null sets (Lebesgue measure), we can define the following equivalence classes:

$$[f] \coloneqq \{g \in \mathcal{L}^2(\mathbb{R}^n) | f \sim g\} \, .$$

Definition E.0.1.
The set of equivalence classes of functions from $\mathcal{L}^2(\mathbb{R}^n)$ gets denoted by $L^2(\mathbb{R}^n)$ and is called $\boldsymbol{L^2}$ **Lebesgue space** . The hermitian scalar product is defined to be

$$\langle f, g \rangle \coloneqq \int_{\mathbb{R}^n} \overline{f(x)} \cdot g(x) \, dx^n \, .$$

E.1. Fourier transformation

Definition E.1.1.
Let $f : \mathbb{R} \to \mathbb{C}$ an integrable function with $\int\limits_{\mathbb{R}} |f(x)| dx < \infty$, i.e. $f \in L^1(\mathbb{R})$. Then:

$$\tilde{f} \coloneqq \mathcal{F} f : \mathbb{R} \longrightarrow \mathbb{C} \, , \qquad k \longmapsto \frac{1}{\sqrt{2\pi}} \int\limits_{\mathbb{R}} f(x) e^{-ikx} dx$$

is called the **Fourier transformed function** of f.

Remark E.1.2.
For functions $f : \mathbb{R}^n \to \mathbb{C}$ the Fourier transformation is defined by

$$\mathcal{F}(f)(\vec{k}) = \frac{1}{\sqrt{2\pi}^n} \int_{\mathbb{R}^n} f(\vec{x}) e^{-i\langle \vec{k}, \vec{x}\rangle} d(x_1, ..., x_n) \ .$$

Definition E.1.3.
The **Schwartz space** $\mathcal{S}(\mathbb{R}^n)$ is the set of rapidly decreasing functions, defined by:

$$\mathcal{S}(\mathbb{R}^n) = \{f \in C^\infty(\mathbb{R}^n) | \forall \alpha, \beta \in \mathbb{N}_0^n : \sup_{x \in \mathbb{R}^n} |x^\alpha \partial^\beta f(x)| < \infty\} \ .$$

Proposition E.1.4 (Inverse Fourier transformation).
For function in $\mathcal{S}(\mathbb{R}^n)$ there exists an inverse Fourier transformation:

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \tilde{f}(k) e^{ikx} dk \ .$$

Remark E.1.5.
It is not uncommon to choose different normalizations than $\frac{1}{\sqrt{2\pi}}$. However, the right pair of normalization constants for transformation and inverse transformation have to be used. In quantum mechanics the following normalization is common:

$$\mathcal{F}(f)(k) = \int_{\mathbb{R}} f(x) e^{-ikx} dx \qquad \text{and} \qquad \mathcal{F}^{-1}(\tilde{f})(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \tilde{f}(k) e^{ikx} dk \ .$$

One reason to study Fourier transformations is the behavior of derivatives under Fourier transformations. This property can be used to solve differential equations. However this does not come for free, as the inverse transformation is needed to get the wanted representation of the solution.

Theorem E.1.6.
*Let P be the **derivative operator**, defined by $(Pf)(x) = -if'(x)$ and Q the **multiplication operator**, defined by $(Qf)(x) = x \cdot f(x)$.*

(i) Let $f \in C^m(\mathbb{R})$ and $P^n f \in L^1(\mathbb{R})$, with $m \geq n$, then:

$$\boxed{\mathcal{F}(P^n f) = Q^n \mathcal{F}(f)}$$

(ii) Let $Q^n f \in L^1(\mathbb{R})$. Then, for $m \geq n$ it follows that$\mathcal{F}(f) \in C^m(\mathbb{R})$ and:

$$\boxed{\mathcal{F}(Q^n f) = (-1)^n P^n \mathcal{F}(f)}$$

Proof E.1.7.
We will only show the initial case $n = 1$. The rest follows from a simple induction.

(i) By assumption it holds that $\lim_{x\to\pm\infty} f'(x) = 0$ and thus $\lim_{x\to\pm\infty} f'(x)e^{-inx} = 0$.[1] Hence:

$$\begin{aligned}\sqrt{2\pi}\ \mathcal{F}(Pf)(k) &= \int_{\mathbb{R}} (Pf)(x)e^{-ikx}dx = \int_{\mathbb{R}} -if'(x)e^{-ikn}dx \\ &= \underbrace{-if(x)e^{-ikx}|_{x=-\infty}^{x=\infty}}_{=0} + \int_{\mathbb{R}} kf(x)e^{-ikx}dx \\ &= k\int_{\mathbb{R}} f(x)e^{-ikx}dx = \sqrt{2\pi}\ Q\mathcal{F}(f)(x)\ .\end{aligned}$$

(ii)

$$\begin{aligned}\sqrt{2\pi}\ \mathcal{F}(Qf)(k) &= \int_{\mathbb{R}} xf(x)e^{-ikx}dx = \int_{\mathbb{R}} -i\frac{d}{dk}f(x)e^{-ikx}dx \\ &= -i\frac{d}{dk}\int_{\mathbb{R}} f(x)e^{-ikx}dx = -P\mathcal{F}(f)(k)\ .\end{aligned}$$

□

Another property of Fourier transformations, that helps calculating the inverse transformation, is the behavior of convolutions.

Definition E.1.8.
Let $f, g : \mathbb{R} \to \mathbb{C}$ be two functions. Then

$$(f * g)(x) := \int_{\mathbb{R}} f(\tau)g(x-\tau)d\tau \underset{\tau=x-t}{=} -\int_{-\mathbb{R}} g(t)f(x-t)dt = (g * f)(x)$$

is called the **convolution** of f and g.

Theorem E.1.9 (Convolution theorem)**.**

$$\mathcal{F}(f * g) = \sqrt{2\pi}\mathcal{F}(f)\cdot\mathcal{F}(g) \qquad \textit{and} \qquad \mathcal{F}(f) * \mathcal{F}(g) = \sqrt{2\pi}\mathcal{F}(f\cdot g)$$

Proof E.1.10.
The proof only uses the commutativity of convolutions and Fubini's theorem, that allows to change the order of integration. □

[1] $|e^{ix}| = 1\ \forall\ x \in \mathbb{R}$

E.2. Physical conventions

Although we do not focus on physical properties here, we will directly call the following quantities appropriately:

Position operator: $\hat{\mathbf{x}} = \vec{x}\,\cdot$

Momentum operator: $\hat{\mathbf{p}} = -i\hbar\nabla = -i\hbar \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix}$

Angular momentum operator: $\hat{\mathbf{L}} = \hat{\mathbf{x}} \times \hat{\mathbf{p}}$

As we have already seen in corollary D.1.18, the angular momentum operator is a representation of the generators of rotations on $L^2(\mathbb{R}^3)$. Comparing to theorem E.1.6, one can see that the position operator is the multiplication operator and the momentum operator is the derivative operator up to a factor $\hbar$.

Definition E.2.1.
Let $|f\rangle \in L^2(\mathbb{R}^3)$ be a Hilbert vector. For a $\vec{x} \in \mathbb{R}^3$ the **position representation** is defined by

$$\langle \vec{x}|f\rangle = f(\vec{x}) \qquad \text{and} \qquad \langle f|\vec{x}\rangle = \overline{f(\vec{x})}\ .$$

For a $\vec{k} \in \mathbb{R}^3$, the **momentum representation** is defined by

$$\langle \vec{p}|f\rangle = (\mathcal{F}f)(\hbar\vec{k}) \qquad \text{and} \qquad \langle f|\vec{p}\rangle = \overline{(\mathcal{F}f)}(\hbar\vec{k})\ .$$

Here $\mathcal{F}$ is the Fourier transformation with quantum mechanical normalization convention $(\mathcal{F}f)(k) = \int_{\mathbb{R}^3} f(\vec{x})e^{-i\vec{k}\cdot\vec{x}}dx^3$.

The name *momentum representation* can be understood from the point of view of wave mechanics and the *De-Broglie relation*. The momentum of a matter wave with wave vector $\vec{k}$ is $\vec{p} = \hbar\vec{k}$.

Definition E.2.2.
Let A be an operator. The notation for the position representation of the operator is

$$\langle \vec{x}|A|f\rangle = Af(\vec{x})\ .$$

This notation becomes more transparent, when one compares the position representation of $\langle g|A|f\rangle$ with the formal definition on $L^2(\mathbb{R}^3)$:

$$\int_{\mathbb{R}^3} \overline{g(\vec{x})}Af(\vec{x})\ dx^3 = \langle g|A|f\rangle = \int_{\mathbb{R}^3} \langle g|\vec{x}\rangle\langle \vec{x}|A|f\rangle\ dx^3\ .$$

Theorem E.1.6 allows to express momentum and position operator in both the position and momentum representation:

	position representation	momentum representation
$\hat{\mathbf{x}}$	$x\cdot$	$i\hbar\nabla_p$
$\hat{\mathbf{p}}$	$-i\hbar\nabla_x$	$p\cdot$

Pseudo eigen vectors and mnemonics

The Dirac notation gives rise to the, at first sight subtle, but conceptually important, question of the meaning of $\langle\vec{x}|$, $\langle\vec{p}|$, $|\vec{x}\rangle$ and $|\vec{p}\rangle$. These objects seem harmless enough. In the physical literature they are even used to define a (pseudo) completeness relation:

$$\int_{\mathbb{R}^3} |\vec{x}\rangle\langle\vec{x}| \; dx^3 = \mathbb{1} = \frac{1}{(2\pi\hbar)^3} \int_{\mathbb{R}^3} |\vec{p}\rangle\langle\vec{p}| \; dp^3 \; .$$

The object $|\vec{x}\rangle$ is supposed to be an eigen state of the position operator $\hat{\mathbf{x}}$, with $\hat{\mathbf{x}}|\vec{x}\rangle = \vec{x}|\vec{x}\rangle$. Similarly the momentum state is tired to be defined by $\hat{\mathbf{p}}|\vec{p}\rangle = \vec{p}|\vec{p}\rangle$.

However, these definitions do not work out mathematically. The objects are no elements of the Hilbert space $L^2(\mathbb{R}^3)$:

$$f(\vec{x}) = \langle\vec{x}|f\rangle = \overline{\langle f|\vec{x}\rangle} = \overline{\int_{\mathbb{R}^3} f(\vec{r})\xi_x(\vec{r}) \; dr^3}$$

$$\Rightarrow \quad |\vec{x}\rangle = \xi_x(\cdot) = \delta(\vec{x} - \cdot) \notin L^2(\mathbb{R}^3) \; ,$$

$$\int_{\mathbb{R}^3} f(\vec{x}) e^{-i\frac{\vec{x}\cdot\vec{p}}{\hbar}} dx^3 = \langle\vec{p}|f\rangle = \int_{\mathbb{R}^3} \overline{\xi_p(\vec{x})} f(\vec{x}) \; dx^3$$

$$\Rightarrow \quad \langle\vec{x}|\vec{p}\rangle = \xi_p(\vec{x}) = e^{i\frac{\vec{x}\cdot\vec{p}}{\hbar}} \quad \Rightarrow \quad |\vec{p}\rangle \notin L^2(\mathbb{R}^3) \; .$$

These mathematical inconsistencies are not only a warning to be careful using Dirac notation, but can serve to understand the meaning of it better. By construction, terms like $f(\vec{x})$ for $f \in L^2(\mathbb{R}^n)$, or in Dirac notation $\langle\vec{x}|f\rangle$, have no meaning on their own. The reason is, that $f \in L^2(\mathbb{R}^n)$ does actually mean the equivalence class $[f]$. Since a single point is a null set, something like $[f](\vec{x})$ has no meaning, it can be chosen to be anything. However, since $\int_U \langle\vec{x}|f\rangle \; dx^3$ is well defined, the term $\langle\vec{x}|f\rangle$ can be understood as density function, that has to be integrated over. Put differently, $\langle\vec{x}|f\rangle$ defines the regular distribution T_f. In that point of view, the term $\langle\vec{x}|\vec{y}\rangle = \delta(\vec{y} - \vec{x})$ can be made sense of, by understanding it as the singular delta distribution $\delta_{\vec{y}}$.[2]

E.3. Spherical harmonics

The position/momentum representation of quantum mechanics are the first representations one encounters. Thus they take a prominent role in solving problems. One important class of problems involves the angular momentum and its eigen states.

[2] There is no such thing as a delta function. Indeed it can be proven, that the delta function is not regular.

E.3.1. Definitions and propositions

Definition E.3.1.
The solution on $[-1,1]$ of the differential equation

$$(1-x^2)\cdot f''(x) - 2x\cdot f'(x) + \ell(\ell+1)\cdot f(x) = 0$$

is called ℓ-th **Legendre polynomial** P_ℓ for all $\ell \in \mathbb{N}_0$. The solutions on $[-1,1]$ of the differential equation

$$(1-x^2)\cdot f''(x) - 2x\cdot f'(x)\left(\ell(\ell+1) - \frac{m^2}{1+x^2}\right) f(x) = 0$$

are called **associated Legendre polynomials** $P_{\ell,m}$ for all $\ell, m \in \mathbb{N}_0$ with $0 < m < \ell$.

Proposition E.3.2.
The following equalities hold:

1. $P_\ell(x) = \dfrac{1}{2^\ell \ell!}\left(\dfrac{d}{dx}\right)^\ell (x^2-1)^\ell$
2. $P_\ell(x) = \dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} x - \sqrt{x^2-1}\cos(\phi)\ d\phi$
   $\Rightarrow \quad P_\ell(\cos(\theta)) = \dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} \cos(\theta) - i\sin(\theta)\cos(\phi)\ d\phi$
3. $\displaystyle\int_{-1}^{1} P_\ell(x)P_{\ell'}(x)\ dx = \frac{2}{2\ell+1}\delta_{\ell\ell'}$
4. $P_\ell(-x) = (-1)^\ell P_\ell(x)$
5. $P_{\ell,m}(x) = (-1)^m(1-x^2)^{\frac{m}{2}}\left(\dfrac{d}{dx}\right)^m P_\ell(x)$

Definition E.3.3.
The **spherical harmonics** $Y_{\ell,m} : [0,\pi]\times[0,2\pi] \to \mathbb{C}$ are functions, that solve the following differential equations simultaneously:

$$\left(\frac{1}{\sin(\theta)}\partial_\theta \sin(\theta)\partial_\theta + \frac{1}{\sin^2(\theta)}\partial_\phi^2\right) Y_{\ell,m} = -\ell(\ell+1)Y_{\ell,m}$$

$$-i\partial_\phi Y_{\ell,m} = mY_{\ell,m}$$

Proposition E.3.4.

The spherical harmonics are orthonormal, i.e.:

$$\int_{S^2} Y_{\ell,m} Y_{\ell',m'} \, d\Omega = \delta_{\ell\ell'} \cdot \delta_{mm'} \ .$$

There is a normalization factor $N_{\ell m} = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{2\ell+1}{2} \frac{(\ell-m)!}{(\ell+m)!}}$*, such that:*

$$Y_{\ell,m}(\theta,\phi) = N_{\ell m} e^{im\phi} P_{\ell m}(\cos(\theta)) \ .$$

E.3.2. "Algebraic" construction of the spherical harmonics

The word "algebraic" is used rather informally here. Usually, to find the spherical harmonics, the differential equations defining them are solved directly, resulting in tedious calculations, giving not much inside at all. Here we want to use a different way to introduce the spherical harmonics, as is done in [Zir16]:

Before doing so, we need the Laplace operator in spherical coordinates:

$$\begin{aligned}
\Delta f &= \star d \star d \ f = \star d \star (\partial_r f \ dr + \partial_\theta f \ d\theta + \partial_\phi f \ d\phi) \\
&= \star d \left(r^2 \sin(\theta) \partial_r f \ [d\theta \wedge d\phi, R] + \sin(\theta) \partial_\theta f \ [d\phi \wedge dr, R] + \frac{\partial_\phi f}{\sin(\theta)} [dr \wedge d\theta, R] \right) \\
&= \star \left(\frac{1}{r^2} \partial_r r^2 \partial_r f + \frac{1}{r^2 \sin(\theta)} \partial_\theta \sin(\theta) \partial_\theta f + \frac{1}{r^2 \sin^2(\theta)} \partial_\phi^2 f \right) dV \\
&= \left(\frac{1}{r^2} \partial_r r^2 \partial_r + \frac{1}{r^2} \left(\frac{1}{\sin(\theta)} \partial_\theta \sin(\theta) \partial_\theta + \frac{1}{\sin^2(\theta)} \partial_\phi^2 \right) \right) f \\
&=: (\Delta_r + \tfrac{1}{r^2} \Delta_{S^2}) f \ .
\end{aligned}$$

In the last line we defined the **spherical Laplace operator** Δ_{S^2}.

We are looking for a differential operator D_ℓ of ℓ-th order, such that

i) $D_\ell \circ \Delta = \Delta \circ D_\ell$

ii) $D_\ell \ \frac{1}{r} = \frac{F_\ell(\frac{x}{r}, \frac{y}{r}, \frac{z}{r})}{r^{\ell+1}}$, where F_ℓ is a polynomial of degree ℓ.

Lemma E.3.5.
The function F_ℓ *is an eigen function of* $-\Delta_{S^2}$ *with eigen value* $\ell(\ell+1)$*.*

Proof E.3.6.

$$0 = \Delta \frac{1}{r} = D_\ell \Delta \frac{1}{r} = \Delta D_\ell \frac{1}{r} = \Delta \frac{F_\ell}{r^{\ell+1}} = \ell(\ell+1) \frac{F_\ell}{r^{\ell+1}} + \Delta_{S^2} \tfrac{F_\ell}{r^{\ell+1}} \ .$$

□

Example E.3.7.
Let $e_u, e_v \in \mathbb{R}^3$ be orthonormal vectors, then $D_\ell = (\mathcal{L}_{e_u+ie_v})^\ell$ is an operator of the kind we are looking for.

Proof E.3.8.
In Cartesian coordinates (u, v, w) the Laplace operator has the form $\Delta = \partial_u^2+\partial_v^2+\partial_w^2$. The Lie derivative acting on functions is given by $\mathcal{L}_v = \iota(v) \circ d$. Thus

$$\mathcal{L}_{e_u+ie_v} = \partial_u + i\partial_v \ .$$

Hence $\mathcal{L}_{e_u+ie_v}$ commutes with Δ and so does $(\mathcal{L}_{e_u+ie_v})^\ell$ commute with Δ.
With $r = \sqrt{u^2 + v^2 + w^2}$ we see that

$$\mathcal{L}_{e_u+ie_v} \frac{1}{r} = \frac{u + iv}{r^2} \ .$$

From $\mathcal{L}_{e_u+ie_v}(u + iv) = 0$, using an iteration, the claim follows. We have also found the form of the polynomial F_ℓ:

$$F_\ell(u, v, w) = (u + iv)^\ell$$

□

Corollary E.3.9.
A special case is the operator $(\mathcal{L}_{e_z+ie_x})^\ell$.

Corollary E.3.10.
Let $D_\ell = (\mathcal{L}_{e_z+ie_x})^\ell$, *then* F_ℓ *has the form*

$$F_\ell\left(\frac{x}{r}, \frac{y}{r}, \frac{z}{r}\right) = \left(\frac{z}{r} + i\frac{x}{r}\right)^\ell = (\cos(\theta) + i\sin(\theta)\cos(\phi))^\ell \ ,$$

and is an eigen function of $-\Delta_{S^2}$ *with the eigen values* $\ell(\ell + 1)$.

The *spherical harmonics* are given by the following expansion

$$(\cos(\theta) + i\sin(\theta)\cos(\phi))^\ell = \sum_{m=-\ell}^{\ell} c_{m,\ell} Y_{m,\ell}(\theta, \phi) \ ,$$

and the following conditions:

$$Y_{\ell,m}(\theta, \phi + \alpha) = e^{im\alpha}\ Y_{\ell,m}(\theta, \phi) \qquad \text{and} \qquad \int_{S^2} |Y_{\ell,m}|^2 \ d\Omega = 1 \ .$$

The first condition can be used to calculate the spherical harmonics:

$$F_\ell(\theta, \phi + \alpha) = \sum_{m=-\ell}^{\ell} e^{im\alpha} c_{m,\ell} Y_{m,\ell}(\theta, \phi) \ .$$

We are now looking for the part of F_ℓ that is proportional to $e^{im\alpha}$ for $\phi \mapsto \phi + \alpha$. This can be achieved by integration:

$$\begin{aligned}\int_0^{2\pi} e^{-im\alpha} F_\ell(\theta, \phi + \alpha)\, d\alpha &= \int_0^{2\pi} e^{-im\alpha} \sum_{m=-\ell}^{\ell} e^{im\alpha} c_{m,\ell} Y_{m,\ell}(\theta, \phi)\, d\alpha \\ &= \beta Y_{m,\ell}(\theta, \phi) \ ,\end{aligned}$$

where the coefficient β has to be determined with the normalization condition, and the convention $\overline{Y_{\ell,m}} = (-1)^m Y_{\ell,-m}$. It remains to show, that these functions solve the defining differential equations:

Theorem E.3.11.
The spherical harmonics satisfy:

$$\begin{aligned}-\Delta Y_{\ell,m} &= \ell(\ell+1) Y_{\ell,m} \ , \\ -i\partial_\phi Y_{\ell,m} &= m Y_{\ell,m} \ , \\ \widehat{P} Y_{\ell,m} &= (-1)^\ell Y_{\ell,m} \ ,\end{aligned}$$

where we define $\widehat{P} f(\vec{r}) = f(-\vec{r})$; ($\widehat{P}$ is called parity operator).

Proof E.3.12.
The function F_ℓ and all relevant derivatives are bounded by integrable functions, such that integration and derivative can be exchanged:

$$-\Delta \beta Y_{\ell,m} = -\Delta \int_0^{2\pi} e^{-im\alpha} F_\ell(\theta, \phi + \alpha)\, d\alpha = \int_0^{2\pi} e^{-im\alpha} - \Delta F_\ell(\theta, \phi + \alpha)\, d\alpha \ .$$

Since $\Delta = (\Delta_r + \frac{1}{r^2}\Delta_{S^2})$ and $r = 1$ on the 2-sphere, using corollary E.3.10, we have proven the first equation.

The second equation follows from

$$\begin{aligned}\partial_\phi Y_{\ell,m} &= \lim_{\alpha \to 0} \frac{Y_{\ell,m}(\theta, \phi + \alpha) - Y_{\ell,m}(\theta, \phi)}{\alpha} = \lim_{\alpha \to 0} \frac{e^{im\alpha} - e^0}{\alpha} Y_{\ell,m}(\theta, \phi) \\ &= \partial_\alpha e^{im\alpha} Y_{\ell,m}(\theta, \phi) = im Y_{\ell,m}(\theta, \phi) \ .\end{aligned}$$

The last equation can easily be seen, using the parity operator in spherical coordinates: $(r, \theta, \phi) \to (r, \pi - \theta, \phi + \pi)$. □

Put in the physical context, the spherical harmonics are eigen functions of the angular momentum operator:

Corollary E.3.13.
The squared angular momentum is $\widehat{L}^2 = -\hbar\Delta$. Using spherical coordinates, one

can show $\widehat{L}_z = -i\hbar\partial_\phi$. *Hence, the spherical harmonics satisfy:*

$$\widehat{L}^2 Y_{\ell,m} = \ell(\ell+1)\hbar^2 Y_{\ell,m} \ ,$$
$$\widehat{L}_z Y_{\ell,m} = m\hbar Y_{\ell,m} \ .$$

F

Distributions and Green's functions

Distributions are the mathematical foundation for many informal calculations in physics. The prime example would be the delta function, that can be found in most physics books. In addition to a more rigorous formulation, distributions allow to solve linear partial differential equations with constant coefficients and give the motivation for Green's functions. Here we will follow [Zir16] and partly [KW06, chapter 12] closely

F.1. Distributions

Distributions are linear functionals on a restricted space of functions, the so called test functions. The meaning of continuity needs the choice of a topology as well, such that some work is needed, before we can give the definition of distributions.

F.1.1. Test functions

In the following we will sometimes use the short hand notation Mit $\int f \, d\lambda$ to denote the (Lebesgue) integration of f. If the functions is also Riemann-integrable it holds that

$$\int f \, d\lambda = \int f(x) \, dx^n \ .$$

Definition F.1.1 (Multi index notation)**.**
A **multi index** is a tupel $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{N}_0^n$. The norm of a multi index is

$$|\alpha| = \alpha_1 + ... + \alpha_n \ .$$

A common usage of multi indices is the short hand notation of partial derivatives:

$$\left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdot ... \cdot \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n} = \partial^\alpha \ .$$

Definition F.1.2.
The **support** $\mathrm{supp}(f)$ of a function $f : \mathbb{R}^n \to \mathbb{C}$, is

$$\mathrm{supp}(f) := \overline{\{x \in \mathbb{R}^n | f(x) \neq 0\}} = \{x \in \mathbb{R}^n | f(x) \neq 0\} \cup \partial\{x \in \mathbb{R}^n | f(x) \neq 0\} \ .$$

Let $\Omega \subset \mathbb{R}^n$ be open and not empty. The set of C^k-functions $f : \Omega \to \mathbb{K}$ witch compact support in Ω commonly is denoted by

$$C_c^k(\Omega) := \{f \in C^k(\Omega) | \mathrm{supp}(f) \subset \Omega \text{ is compact}\} \ .$$

Lemma F.1.3.
Let $f, g \in C_c^k(X)$, where $X \subset \mathbb{R}^n$ can be written as direct product, and let $\alpha \in \mathbb{N}^n$ be a multi index with $|\alpha| \leq k$, then:

$$\int_X (\partial^\alpha f) \cdot g \; d\lambda = (-1)^{|\alpha|} \int_X f \cdot \partial^\alpha g \; d\lambda \; .$$

Proof F.1.4.
It is enough to use partial integration and Fubini's theorem. Let Œ $\partial^\alpha = \partial^{\alpha'} \circ \partial_{x_1}$, i.e. $|\alpha'| = |\alpha| - 1$,. By assumption $X = X' \cup I$:

$$\begin{aligned}\int_X \partial^\alpha f \cdot g \; d\lambda &= \int_{X'} \left(\int_I \partial^{\alpha'} \partial_{x_1} f(x_1, ..., x_n) \cdot g(x_1, ..., x_n) \; dx_1 \right) dx_2 ... dx_n \\ &= \int_{X'} \Big(\partial^{\alpha'} f(x_1, ..., x_n) \cdot g(x_1, ..., x_n) \Big|_{\partial I} \\ &\qquad - \int_I \partial^{\alpha'} f(x_1, ..., x_n) \cdot \partial_{x_1} g(x_1, ..., x_n) \; dx_1 \Bigg) dx_2 ... dx_n\end{aligned}$$

The functions have compact support in X, hence:

$$\partial^{\alpha'} f(x_1, ..., x_n) \cdot g(x_1, ..., x_n) \Big|_{\partial I} = 0$$

and thus:

$$\begin{aligned}&\int_{X'} \left(\int_I \partial^{\alpha'} \partial_{x_1} f(x_1, ..., x_n) \cdot g(x_1, ..., x_n) \; dx_1 \right) dx_2 ... dx_n \\ &= - \int_{X'} \left(\int_I \partial^{\alpha'} f(x_1, ..., x_n) \cdot \partial_{x_1} g(x_1, ..., x_n) \; dx_1 \right) dx_2 ... dx_n\end{aligned}$$

Repeating the above steps proves the claim. □

Definition F.1.5.
C^∞-functions with compact support in Ω, $f \in C_c^\infty(\Omega)$, are called **test functions**. The set of test functions is denoted by $\mathcal{D}(\Omega) \equiv C_c^\infty(\Omega)$.

Lemma F.1.6.
The function $f : \mathbb{R} \to [0, \infty), \; t \mapsto f(t)$ with

$$f(t) = \begin{cases} e^{\frac{-1}{t}} & t > 0 \\ 0 & t \leq 0 \end{cases}$$

is smooth on $\mathbb{R}$.

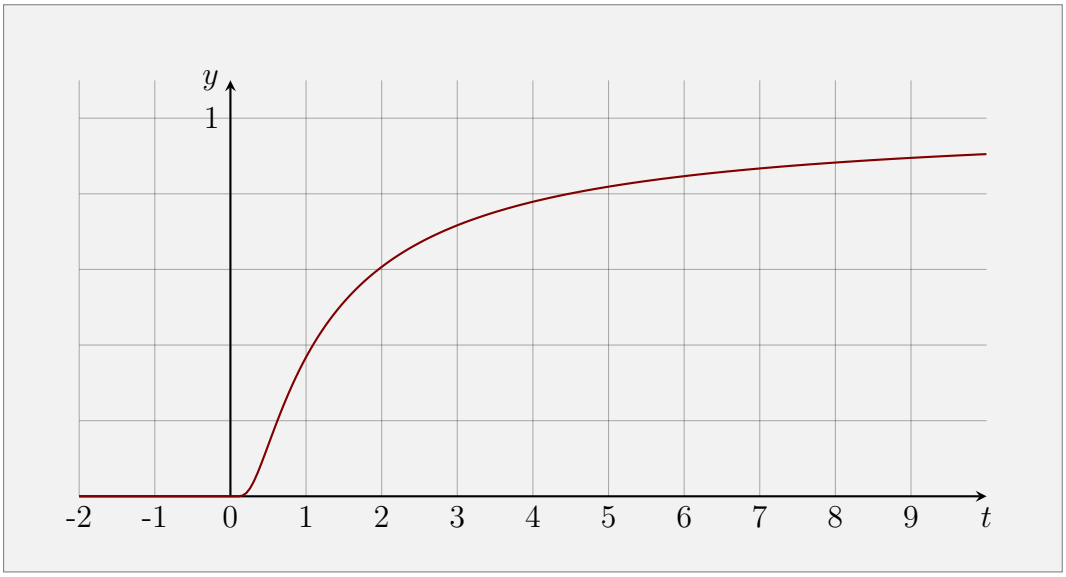


Figure F.1.: Plot of $f(t)$ from lemma F.1.6

Proof F.1.7.

1. For $t \neq 0$ it is a composition of C^∞-functions, and thus smooth.
2. It is sufficient to show $\lim\limits_{t\to 0} \frac{e^{-\frac{1}{t}}}{t^n} = 0$:

$$\lim_{t\to 0}\left|\frac{e^{-\frac{1}{t}}}{t^n}\right| = \lim_{x\to\infty}\left|\frac{e^{-x}}{x^{-n}}\right| = \lim_{x\to\infty}\left|\frac{x^n}{e^x}\right| \geq \lim_{x\to\infty}\left|\frac{x^n}{x^{+1}}\right| = 0 \ .$$

□

Theorem F.1.8.
*For all $\varepsilon > 0$ there is a test function $j_\varepsilon \in \mathcal{D}(\Omega)$, called **bump function**, such that:*

$$j_\varepsilon \geq 0 \ , \quad \operatorname{supp}(j_\varepsilon) = \overline{B_\varepsilon(0)} \quad \textit{and} \quad \int_\Omega j_\varepsilon \, d\lambda = 1 \ .$$

Proof F.1.9.
Let $\psi_\varepsilon : \mathbb{R}^n \to [0,\infty), x \mapsto \psi_\varepsilon(x)$ be defined by

$$\psi_\varepsilon(x) = \begin{cases} e^{\frac{-1}{1-||x||^2/\varepsilon^2}} & ||x|| < \varepsilon \\ 0 & \text{else} \end{cases} \ .$$

By construction: $\operatorname{supp}(\psi_\varepsilon) \subseteq B_\varepsilon(0)$. The C^∞-differentiability follows from lemma F.1.6. Thus $\psi_\varepsilon \in \mathcal{D}(\Omega)$. Using the following normalization, the last property of j_ε can be realized:

$$j_\varepsilon = \frac{\psi_\varepsilon}{\int\limits_\Omega \psi_\varepsilon \, d\lambda} \ .$$

□

Definition F.1.10.
The **convolution** of functions $f, g : \mathbb{R}^n \to \mathbb{K}$ is defined by:

$$(f * g)(x) = \int_{\mathbb{R}^n} f(y) \cdot g(x - y)\ dy^n\ .$$

Lemma F.1.11.
For $f \in C_c^k(\mathbb{R}^n)$ and $g \in \mathcal{L}_{lok}(\mathbb{R}^n)$ it holds that:

i) The convolution is k-times differentiable $f * g \in C^k(\mathbb{R}^n)$.

ii) The first k partial derivatives commute with the convolution integral:

$$\partial^\alpha (f * g) = (\partial^\alpha f) * g \quad f''ur\ |\alpha| \leq k\ .$$

iii) $\operatorname{supp}(f * g) = \operatorname{supp}(f) + \operatorname{supp}(g) := \{x + y \in \mathbb{R}^n | x \in \operatorname{supp}(f),\ y \in \operatorname{supp}(g)\}.$

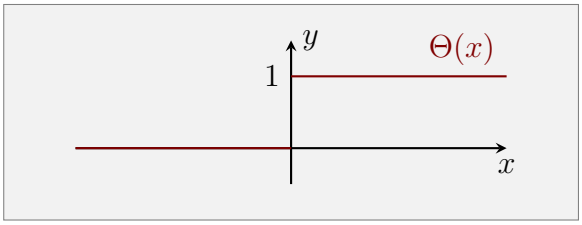
iv) Let $A = \operatorname{supp}(g)$, then:

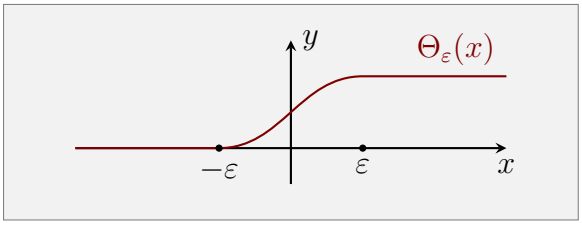
$$j_\varepsilon * g \in C^\infty(\mathbb{R}^n) \quad \text{and} \quad \operatorname{supp}(j_\varepsilon * g) \subset \bigcup_{a \in A} \overline{B_\varepsilon(a)}\ .$$

Definition F.1.12.
Let $f : \Omega \subset \mathbb{R}^n \to \mathbb{K}$ be local inetgrabl and set $f(x) = 0\ \forall\ x \in \mathbb{R}^n \setminus \Omega$. The set of functions $f_\varepsilon = f * j_\varepsilon$ is called **smoothing** of f.

Example F.1.13 (Heaviside function)**.**
The Heaviside function is $\Theta : \mathbb{R} \to \{0, 1\},\ x \mapsto \chi_{[0,\infty)}(x)$.
The smoothing of $\Theta_\varepsilon = \Theta * j_\varepsilon$ strictly monotonically increasing on $[-\varepsilon, \varepsilon]$, and $\Theta_\varepsilon(x) = 0$ for $x \leq -\varepsilon$ and $\Theta_\varepsilon(x) = 1$ for $x \geq \varepsilon$. These properties follow from the convolution integral:

$$\Theta_\varepsilon(x) = \int_{\mathbb{R}} \Theta(x - y) j_\varepsilon(y) dy = \int_{-\varepsilon}^{x} j_\varepsilon(y)\ dy$$





Theorem F.1.14.

i) Let $f \in L^1_{loc}(\Omega)$ and $f_\varepsilon = f * j_\varepsilon$, then, for every compact subset $K \subset \Omega$:

$$\lim_{\varepsilon \to 0} \int_K |f - f_\varepsilon| \, d\lambda = 0 \ .$$

ii) Let Ω be open and $f \in C^0(\Omega)$ with $\operatorname{supp}(f) \subset \Omega$, then:

$$\exists\, \varepsilon > 0 \ \forall\, r < \varepsilon : \ f_r = f * j_r \ \in \mathcal{D}(\Omega) \quad \text{and} \quad f_r \overset{r \to 0}{\longrightarrow} f \ \text{uniformly in } \Omega \ .$$

Corollary F.1.15.

- *The set test functions is dense in $L^1_{lok}(\Omega)$ with respect to the L^1-norm.*
- *The set of test functions is dense in $C^0_c(\Omega)$ w.r.t. the uniform norm.*

F.1.2. Distributions

To define distributions a stronger sense of convergence is needed on $\mathcal{D} := C^\infty_c(\mathbb{R}^n)$.

Definition F.1.16.
Let (f_n) be a sequence of test functions $f_n \in \mathcal{D}$. This sequence is called **convergent on** $\mathbb{D}$test function! convergence with limit $f \in \mathcal{D}$, ief:

i) $\exists\ K \subset \mathbb{R}^n$ compact, such that $\operatorname{supp}(f_n) \subset K\ \forall\ n \in \mathbb{N}$.

ii) $\forall\ \alpha \in \mathbb{N}^n : \ \partial^\alpha f_n \longrightarrow \partial^\alpha f$ uniformly for $n \to \infty$.

In this case we write $f_n \overset{\mathcal{D}}{\to} f\ n \to \infty$.

Definition F.1.17.
A continuous linear map $T : \mathcal{D} \to \mathbb{C}$ is called **distribution** . Continuity means

$$f_n \overset{\mathcal{D}}{\to} f \quad \Rightarrow \quad T[f_n] \overset{n \to \infty}{\longrightarrow} T[f] \ .$$

The vector space of distributions is denoted by $\mathcal{D}\,'$.

Theorem F.1.18.
Let $f : \mathbb{R}^n \to \mathbb{C}$ with $f \in L^1_{lok}(\mathbb{R}^n)$. The map $T_f : \mathcal{D} \to \mathbb{C}$ defined by

$$T_f[\varphi] = \int_{\mathbb{R}^n} f \cdot \varphi \, d\lambda \quad \text{für } \varphi \in \mathcal{D}$$

is a distribution.

Proof F.1.19.
Let (φ_n) be a convergent sequence in $\mathcal{D}$. That means, there is a compact set $K \subset \mathbb{R}^n$ with $\text{supp}(f_n) \subset K$, $\forall\, n \in \mathbb{N}$, such that the integrals exist. It follows:

$$|T_f[\varphi_n] - T_f[\varphi]| \overset{def.}{=} \left| \int_{\mathbb{R}^n} f \cdot (\varphi_n - \varphi) \, d\lambda \right| \leq \max_{x \in K} |\varphi_n(x) - \varphi(x)| \cdot \int_K f \, d\lambda \overset{n\to\infty}{\longrightarrow} 0 \, .$$

Linearity is a property of the integral. □

Definition F.1.20.
A distribution that can be defined by a function f with the help of integration $T = T_f$ is called **regular distribution**. Otherwise the distribution is called **singular**.

Definition F.1.21.
Let (T_n) be a sequence of distributions in $\mathcal{D}$. This sequence is called **convergent in $\mathcal{D}'$** with limit T, if:

$$T_n[\varphi] \overset{n\to\infty}{\longrightarrow} T[\varphi] \; \forall\, \varphi \in \mathcal{D} \, .$$

In that case we write $T_n \overset{\mathcal{D}'}{\longrightarrow} T$.

Considering a function $f \in C^1(\mathbb{R}^n)$, then

$$(T_{\partial_i f})[\varphi] = \int_{\mathbb{R}^n} \partial_i f \cdot \varphi \, d\lambda = -\int_{\mathbb{R}^n} f \cdot \partial_i \varphi \, d\lambda = -T_f[\partial_i \varphi]$$

because of the compact support. This motivates the following definition:

Definition F.1.22.
Let $T \in \mathcal{D}'$ be a distribution and $\alpha \in \mathbb{N}^n$ a multi index. The partial derivative $\partial^\alpha T$ is defined by:

$$(\partial^\alpha T)[\varphi] = (-1)^{|\alpha|} \, T[\partial^\alpha \varphi] \, .$$

Understanding functions as regular distributions, the previous definition allows to differentiate L^1_{loc}-functions. However, since these functions are not differentiable in the usual sense, these derivatives are called **weak derivative**.

Definition F.1.23.
Let $f \in C^\infty(\mathbb{R}^n)$ and $T \in \mathcal{D}'$, then $f \cdot T$ is defined by

$$(f \cdot T)[\varphi] = T[f \cdot \varphi] \ .$$

Corollary F.1.24.
Weak derivatives satisfy the product rule: $\partial_i(f \cdot T) = (\partial_i f) \cdot T + f \cdot \partial_i T$.

F.1.3. Delta distribution

A very important distribution is the **Delta distribution** $\delta_x \in \mathcal{D}'$, defined by

$$\delta_x[f] = f(x) \ .$$

In physics books one often finds Delta functions, however, the following theorem shows that no such thing exists:

Theorem F.1.25.
The Delta distribution δ_x *is singular.*

Proof F.1.26 (Œ for $\boldsymbol{x = 0}$ and $\mathbb{R}$)**.**
Suppose there is a function $\delta(x)$, such that $T_\delta = \delta_0$. Then

$$\int_{\mathbb{R}} \delta(x) \cdot \varphi(x) \ dx = \varphi(0) \quad \forall \ \varphi \in \mathcal{D} \ .$$

Consider the function

$$\varphi_b(x) = \begin{cases} e^{\frac{b^2}{x^2 - b^2}} & , \ |x| < b \\ 0 & , \ |x| \geq b \end{cases} \qquad \Rightarrow \quad \delta_0[\varphi_b] = e^{-1} \ .$$

An estimate of the integration results in a contradiction:

$$\varphi_b(0) = |\varphi_b(0)| = \left| \int_{\mathbb{R}} \delta(x) \cdot \varphi_b(x) \ dx \right| = \left| \int_{-b}^{b} \delta(x) \cdot \varphi_b(x) \ dx \right|$$

$$\leq \underbrace{\sup_{x \in [-b,b]} |\varphi_b(x)|}_{= \varphi_b(0)} \cdot \int_{-b}^{b} |\delta(x)| \ dx \ .$$

For b small enough, it holds that $\int_{-b}^{b} |\delta(x)| \ dx < 1$, and thus $\varphi_b(0) \leq c \cdot \varphi_b(0)$ for $c < 1$. □

Although there is no Delta function, there are sequences of functions (δ_n), called **Dirac sequences**, such that

$$\lim_{n \to \infty} \int_{\mathbb{R}} \delta_n(x) \cdot \varphi(x) \ dx = \varphi(0) \ .$$

Example F.1.27.
As Dirac sequence we use rectangle functions:

$$t_n : [-1,1] \longrightarrow \mathbb{R},\ x \longmapsto \begin{cases} \frac{n}{2} & ,\ x \in [-\frac{1}{n}, \frac{1}{n}] \\ 0 & ,\ \text{sonst} \end{cases} .$$

With the mean value theorem, that says

$$\forall\ f \in C^0([a,b])\ \exists\ \xi \in [a,b] :\ \int_a^b f(x)\ dx = f(\xi) \cdot (b-a)\ ,$$

we find:

$$\forall\ n \in \mathbb{N}\ \exists \xi_n \in \left[-\frac{1}{n}, \frac{1}{n}\right] :\ T_{t_n}[\varphi] = \int_{-1/n}^{1/n} \frac{n}{2}\varphi(x)\ dx = \frac{n}{2}\varphi(\xi_n) \cdot \frac{2}{n} \overset{n\to\infty}{\longrightarrow} \varphi(0)\ .$$

Hence:

$$\lim_{n\to\infty} T_{t_n} = \delta_0\ .$$

F.2. Fundamental solutions and Green's functions

Distributions can be used to solve linear partial differential equations. A central concept therefore are fundamental solutions, that are closely related to Green's functions.

Definition F.2.1.
Let $\alpha \in \mathbb{N}^n$ be a multi index and $c_\alpha \in C^0(\mathbb{R}^n)$ be a continuous function. Then

$$L = \sum_{|\alpha| \le m} c_\alpha \cdot \partial^\alpha$$

is called a **linear differential operator** of order m on $\mathbb{R}^n$.

Definition F.2.2.
The **formally adjoint operator** L^* of L is defined by

$$L^* f = \sum_{|\alpha| \le m} (-1)^{|\alpha|} \partial^\alpha (c_\alpha \cdot f)\ .$$

Lemma F.2.3.
Let $\varphi \in C_c^m(\mathbb{R}^n)$ and $f \in C^m(\mathbb{R}^n)$, then for every linear differential operator L of

order m, it holds that:

$$\int_{\mathbb{R}^n} fL^*\varphi \, d\lambda = \int_{\mathbb{R}^n} Lf \cdot \varphi \, d\lambda \, .$$

Proof F.2.4.
All we need is lemma F.1.3:

$$\begin{aligned}\int_{\mathbb{R}^n} fL^*\varphi \, d\lambda &= \sum_{|\alpha|\leq m} \left((-1)^{|\alpha|} \int_{\mathbb{R}^n} f\partial^\alpha (c_\alpha \cdot \varphi) \, d\lambda \right) \\ &= \sum_{|\alpha|\leq m} \left(\int_{\mathbb{R}^n} \partial^\alpha f \cdot (c_\alpha \cdot \varphi) \, d\lambda \right) \\ &= \sum_{|\alpha|\leq m} \left(\int_{\mathbb{R}^n} c_\alpha \partial^\alpha f \cdot \varphi \, d\lambda \right) = \int_{\mathbb{R}^n} Lf \cdot \varphi \, d\lambda \, .\end{aligned}$$

□

Definition F.2.5.
Let L be a linear differential operator of order m. A function $u \in L^1_{loc}(\Omega)$ is called **weak solution** with inhomogenity $f \in L^1_{loc}(\Omega)$ of the differential equation $Lu = f$, if:

$$\int_\Omega u \, L^*\varphi \, d\lambda = \int_\Omega f \cdot \varphi \, d\lambda \quad \forall \, \varphi \in \mathcal{D} \, .$$

Theorem F.2.6.
Let $Lu = f$ be a differential equation with $f \in C^0(\Omega)$, then:

A function $u \in C^m(\Omega)$ is a weak solution $\Leftrightarrow$ u is a classical solution.

Proof F.2.7.
If $Lu = f$ is satisfied, then $\int \Omega Lu \cdot \varphi - f \cdot \varphi \, d\lambda = 0 \; \forall \; \varphi \in \mathcal{D}$. The reverse direction is a consequence of the fundamental lemma of variational calculus. The rest follows from lemma F.2.3. □

Corollary F.2.8.
Let $u, f \in L^1_{loc}(\mathbb{R}^n)$ and let L be a liner differential operator. Considering the differential equation $Lu = f$ together with the definition of regular distributions T_u, we find

$$LT_u = T_f \quad \Leftrightarrow \quad u \text{ is a weak solution of } Lu = f \, .$$

This corollary motivates to generalize differential equations to distributions:

Definition F.2.9.
Let L be a linear differential operator. A distribution $U \in \mathcal{D}'$ is called **fundamental solution** for L with pole $x \in \mathbb{R}^n$, if:

$$LU = \delta_x \ .$$

Fundamental solutions are not unique. If T is a solution of $LT = 0$, and U is a fundamental solutions, so is $T + U$.

Definition F.2.10.
Let $T \in \mathcal{D}'$ be a distribution and $\varphi \in \mathcal{D}$ a test function. The **convolution** $T * \varphi$ is defined by

$$(T * \varphi)(x) = T[\varphi \circ \tau_x] \ ,$$

with $\tau_x(y) = x - y$.

Lemma F.2.11.
It holds that: $\delta_0 * \varphi = \varphi$

Theorem F.2.12.
Let $U \in \mathcal{D}'$ be a distribution and $\varphi \in \mathcal{D}$. Then the function $u = T * \varphi$ is in $C^\infty(\mathbb{R}^n)$ and

$$\partial_i u = (\partial_i T) * \varphi \ .$$

Proof F.2.13 (Œ in one dimension)**.**

$$\begin{aligned} u'(x) &= \lim_{h\to 0} \frac{u(x+h) - u(x)}{h} = \lim_{h\to 0} \frac{T[\varphi \circ \tau_{x+h}] - T[\varphi \circ \tau_x]}{h} \\ &= \lim_{h\to 0} T\left[\frac{\varphi \circ \tau_{x+h} - \varphi \circ \tau_x}{h}\right] = T[-\varphi' \circ \tau_x] = T'[\varphi \circ \tau_x] = T' * \varphi \end{aligned}$$

□

Corollary F.2.14.
Let L be a linear differential operator with constant coefficients, then:

$$L(T * \varphi) = (LT) * \varphi$$

The reason to consider fundamental solutions is the following theorem:

Theorem F.2.15.
Let L be a linear differential operator with constant coefficients and $T \in \mathcal{D}'$ be a fundamental solution with pole in 0.

i) Let $\varphi \in \mathcal{D}$ be a test function, then $u := T * \varphi$ is a solution of

$$Lu = \varphi \ .$$

ii) If T is a regular distribution to $g \in L^1_{loc}\mathbb{R}^n$, i.e. $T = T_g$, then $u = g * f$ is a weak solution of

$$Lu = f \ .$$

Proof F.2.16.

1.
$$L(T * \varphi) = (LT) * \varphi = \delta_0 * \varphi = \varphi \ .$$

2. Because of F.2.8 it is enough to show, that $LT_u = T_f$ holds. Using Fubini's theorem:

$$\begin{aligned}
(LT_u)[\varphi] = T_u[L^*\varphi] &= \int\limits_{\mathbb{R}^n}\int\limits_{\mathbb{R}^n} g(y-x)f(x)dx^n \ L^*\varphi(y) \ dy^n \\
&= \int\limits_{\mathbb{R}^n}\int\limits_{\mathbb{R}^n} g(y-x)L^*\varphi(y) \ dy^n \ f(x)dx^n \\
&= \int\limits_{\mathbb{R}^n} T_g[L^*(\varphi \circ \tau_y)](x) \ f(x)dx^n \\
&= \int\limits_{\mathbb{R}^n} (LT_g * \varphi)(x)f(x) \ dx^n = \int\limits_{\mathbb{R}^n} (\delta_0 * \varphi)(x)f(x) \ dx^n \\
&= \int\limits_{\mathbb{R}^n} f(x)\varphi(x) \ dx^n = T_f[\varphi] \ .
\end{aligned}$$

□

Definition F.2.17.
A **Green's function** $G(y,x)$ for a linear differential operator L is a fundamental solution with pole in y: $LT_G = \delta_y$, where T_G is the regular distribution w.r.t. $x \mapsto G(y,x)$.

Using the delta function notation of physics literature, Green's functions have to satisfy $LG(y,x) = \delta(x-y)$.

Definition F.2.18.
We define the **shift** of a function to be

$$v_y f(x) = f(y-x) \ .$$

For a distribution we set

$$v_y U[\varphi] = U[v_{-y}\varphi] \ .$$

Lemma F.2.19.
Let U be a fundametal solution with pole in y, for an L with constant coefficients. Then $v_{-y}U$ is a fundamental solution with pole in 0.

Proof F.2.20.

$$L(v_{-y}U)[\varphi] = U[v_y L^*\varphi] = [L^* v_y\varphi] = LU[v_y\varphi] = \delta_y[v_y\varphi] = \varphi(y-y) = \varphi(0) \ .$$

□

Corollary F.2.21.
Let $G(y,x)$ be a Green's function, then $g(x) = G(y, y-x)$ is a fundamental solution with pole in 0. That is: $g(y-x) = G(y,x)$. The solution of the differential equation $Lu = f$ is $u(y) = (g * f)(y)$. Thus the (weak) Green's function has the following defining property:

$$u(y) = \int\limits_{\mathbb{R}^n} g(y-x)f(x)dx^n = \int\limits_{\mathbb{R}^n} G(y,x)f(x)dx^n \ .$$

Proof F.2.22.
All that has to be shown here is, that $v_{-y}T_G = T_{G'}$ with $G'(y,x) = G(y, y-x)$:

$$v_{-y}T_G[\varphi] = T_G[v_y\varphi] = \int_{\mathbb{R}^n} G(y,x)\varphi(y-x)dx^n \overset{x=y-z}{=}$$
$$\int_{\mathbb{R}^n} G(y, y-z)\varphi(z)dz^n = T_{G'}[\varphi] \ .$$

□

F.3. Fourier transformation of tempered distributions

Using the Schwartz space together with a stronger sense of convergence allows to define a tempered distributions, that act on a larger class of functions:

$$\varphi_k \xrightarrow{\mathcal{S}} \varphi \ , \ k \to \infty \qquad \Leftrightarrow \quad x^\alpha \partial^\beta \varphi_k \to \varphi \quad \forall \ \alpha, \beta \in \mathbb{N}_0^n \ .$$

Definition F.3.1.
A **tempered distribution** T on $\mathbb{R}^n$ is a continuous linear map $\mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$. The vector space of tempered distributions is denoted by $\mathcal{S}\,'(\mathbb{R}^n)$.

$$\begin{array}{ll} \text{test functions:} & C_c^\infty(\mathbb{R}^n) \equiv \mathcal{D}(\mathbb{R}^n) \ \subset \ \mathcal{S}(\mathcal{R}^n) \ \subset \mathcal{C}(\mathbb{R}^n) \equiv C^\infty(\mathbb{R}^n) \\ \text{distributions:} & \mathcal{D}'(\mathbb{R}^n) \ \supset \mathcal{S}'(\mathbb{R}^n) \ \supset \mathcal{C}'(\mathbb{R}^n) \end{array}$$

It can be shown, that distributions with compact support are tempered distributions. The motivation to define this kind of distribution is the Fourier transformation:

Definition F.3.2.
The **Fourier transformation** $\mathcal{F}T$ of a tempered distribution T is defined by:

$$\mathcal{F}T[\varphi] = T[\mathcal{F}\varphi] \ .$$

Theorem F.3.3.

i) The Fourier transformation of a tempered distribution is again tempered.

ii) The Fourier transformation $\mathcal{F} : \mathcal{S}\,' \to \mathcal{S}\,'$ is a bijection.

iii) The multiplication and derivatives behave as usual:

$$\mathcal{F}(\partial^\alpha T) = i^{|\alpha|} k^\alpha \mathcal{F}T \qquad \text{and} \qquad \mathcal{F}(k^\alpha T) = (-1)^{|\alpha|} \cdot i^{|\alpha|+1} \mathcal{F}T \ .$$

A common usage of tempered distributions in quantum mechanics is the connection between Delta distributions and exponential functions $e_a(x) \coloneqq e^{ixa}$:

Lemma F.3.4.

$$\mathcal{F}\delta_a = \tfrac{1}{(\sqrt{2\pi})^n} \cdot T_{e_{-a}} \quad \text{and} \quad \tfrac{1}{(\sqrt{2\pi})^n} \cdot \mathcal{F}T_{e_a} = \delta_a \ .$$

Proof F.3.5.

$$\begin{aligned} \mathcal{F}\delta_a[\varphi] = \delta_a[\mathcal{F}\varphi] = (\mathcal{F}\varphi)(a) &= \tfrac{1}{(\sqrt{2\pi})^n} \int_{\mathbb{R}^n} e^{-iax} \cdot \varphi(x) \ dx^n \\ &= \left(\tfrac{1}{(\sqrt{2\pi})^n} \cdot T_{e_{-a}}\right)[\varphi] \ , \end{aligned}$$

$$\begin{aligned} \mathcal{F}T_{e_a}[\varphi] = T_{e_a}[\mathcal{F}\varphi] = \int_{\mathbb{R}^n} e^{iak} (\mathcal{F}\varphi)(k) \ dk^n &= (\sqrt{2\pi})^n \mathcal{F}^{-1}(\mathcal{F}\varphi)(a) \\ = (\sqrt{2\pi})^n \varphi(a) = (\sqrt{2\pi})^n \delta_a[\varphi] \ . \end{aligned}$$

□

Adopting the notation of physics textbooks, that is writing also singular distributions with integral kernel, the delta distribution has the formal representation as delta function, $\delta_a \rightsquigarrow \delta(\bullet - a)$. Thus the action of the delta distribution can be written (only symbolically) as integral

$$\delta_a[\varphi] = \int_{\mathbb{R}^n} dx^n \delta(x - a)\varphi(x) \ .$$

Converting the last lemma to this notation, we get

$$f(x) = e^{iqx} \xrightarrow{(\sqrt{2\pi})^n \mathcal{F}} \widehat{f}(k) = \delta(k-q) \ .$$

Using the lemma in this notation leads to some kind of orthonormality normalization of the integral kernels e^{iqx} and e^{-ipx}:

$$\frac{1}{(\sqrt{2\pi})^n} \int_{\mathbb{R}^n} e^{iqx} e^{-ipx} \ dx^n = \frac{1}{(\sqrt{2\pi})^n} \int_{\mathbb{R}^n} e^{i(q-p)x} \ dx^n = \delta(p-q) \ .$$

Proof F.3.6.
Using the integration kernel notation, we find:

$$\begin{aligned} \mathcal{F}T_{e_q}[\varphi] &= \frac{1}{(\sqrt{2\pi})^n} \int_{\mathbb{R}^n} e^{iqx} \int_{\mathbb{R}^n} e^{-ipx} \varphi(p) \ dp^n \ dx^n \\ &= \int_{\mathbb{R}^n} \frac{1}{(\sqrt{2\pi})^n} \int_{\mathbb{R}^n} e^{iqx} e^{-ipx} dx^n \varphi(p) \ dp^n \\ &= \delta_q[\varphi] = \int_{\mathbb{R}}^n \delta(p-q)\varphi(p) \ dp^n \ . \end{aligned}$$

□

It should be mentioned, that the last equation is only symbolical, meaning that $\delta(p-q)$ needs to be understood as tempered distribution, that has to be integrated against. We might be in some trouble, since $\delta(p-q) = \delta(q-p)$, and it could be suggested to integrate over q instead of p. However, formally we also have

$$\frac{1}{(\sqrt{2\pi})^n} \cdot \mathcal{F}^{-1} T_{e_{-q}} = \delta_q \qquad \rightsquigarrow \qquad \frac{1}{(\sqrt{2\pi})^n} \cdot \frac{1}{(\sqrt{2\pi})^n} \int_{\mathbb{R}^n} e^{i(q-p)x} \ dx^n = \delta(q-p) \ .$$

G
Variational calculus for fields

In this chapter we cover the fundamentals of variational calculus. We introduce the mathematical notion of functionals and the concepts, as well as properties, of their derivatives, following [Wer11, chapter III.5]. In the second section we focus on the important class of functionals that can be expressed by Lagrange-densities, deriving the Euler-Lagrange-equations. We conclude this chapter by discussing notations of the physical literature, used in [AS10] for example.

G.1. Functional derivative and variation

In variational calculus (e.g. in the context of least action) the goal is to find functions, that minimize or maximize a given functional. In general a **functional** is a map from a normed space X into its number field. Recalling, that for functions $f\colon \mathbb{R} \to \mathbb{R}$ a vanishing first derivative is a necessary condition for extrema, we are looking for a similar computational tool in the case of functionals. We will see, that there is a concept of derivative, that is not only connected to the concept of variations, but also the computational tool we are looking for.

Definition G.1.1.
Let $F\colon U(x_0) \subset X \to \mathbb{R}$ be a functional defined on a neighborhood containing x_0. The ***n*-th variation** $\delta^n F(x_0; h)$ of F in direction of $h \in X$ is defined, if it exists, by

$$\delta^n F(x_0; h) = \frac{d^n}{dt^n} F(x_0 + th)\bigg|_{t=0} .$$

The variation resembles the directional derivative of multivariable calculus. This leads to the following definition:

Definition G.1.2.
Let F be a functional as before. If $\delta F(x_0; h)$ exists for all $h \in X$ and there is a linear continuous functional $L\colon X \to \mathbb{R}$, such that

$$L(h) = \delta F(x_0; h) \quad \forall\, h \in X \ ,$$

the functional is called **Gâteaux-differentiable** in x_0. The linear continuous functional is called the **Gâteaux-derivative** $\widehat{G}_{x_0} F$ of F in x_0. Accordingly F is called Gâteaux-differentiable on X if it is Gâteaux-differentiable in all $x \in X$.

As in multivariable calculus, we are not only interested in the directional derivative, but also for the derivative itself. Even in the finite-dimensional case, the existence of the directional derivative in every direction is not enough to prove the existence of the total

derivative. Writing the definition of the Gâteaux-derivative slightly different reveals the generalization necessary to define an analogue to the total differential:

$$\lim_{t\to 0} \frac{|F(x_0 + th) - F(x_0) - t\widehat{G}_{x_0}F(h)|}{t} = 0 \ .$$

In this form one can see easily the similarity to the directional derivative. In the finite-dimensional case one demands the sequence to be uniformly convergent to define the total differential, motivating the following definition:

Definition G.1.3.
Let $F\colon U(x_0) \subset X \to \mathbb{R}$ be a functional on a neighborhood around x_0. The functional is said to be **Fréchet-differentiable** in x_0, if there is a continuous linear functional $L\colon X \to \mathbb{R}$ such that

$$\lim_{X\ni h\to 0} \frac{|F(x_0 + h) - F(x_0) - L(h)|}{\|h\|_X} = 0 \ .$$

The continuous linear functional L is called the **Fréchet-derivative** $D_{x_0}F$ in x_0 of F.

From the definitions it is clear, that if a functional is Fréchet-differentiable in x_0 is is also Gâteaux-differentiable in x_0 and

$$D_{x_0}F(h) = \widehat{G}_{x_0}F(h) = \delta F(x_0; h) \ .$$

The condition to be Fréchet-differentiable can be reinterpreted as linearization of the functional. The existence of the limit is equivalent to

$$F(x_0 + h) - F(x_0) - D_{x_0}F(h) = r(h) \qquad \text{with} \quad \lim_{X\ni h\to 0} \frac{r(h)}{\|h\|_X} = 0 \ .$$

In introductory texts about analytical mechanics the concept of **functional derivative** is introduced as follows: If the change of the functional $F(x_0 + h) - F(x_0)$ can be written as sum of a part $L(x_0, h)$ that is linear in h and a part $R(x_0, h)$ that decreases faster than $\|h\|_X$, i.e.

$$F(x_0 + h) - F(x_0) = L(x_0, h) + R(x_0, h) = D_{x_0}F(h) + r(h) \ ,$$

the functional L is called functional derivative of F in x_0. With the knowledge about Fréchet-derivatives, one can clearly see, that theses texts introduce the Fréchet-derivative, which can be defined in even more general cases. The Fréchet-derivative of an operator $F\colon X \to Y$ can be defined as before, using $\|\cdot\|_Y$ instead of $|\cdot|$.

Theorem G.1.4 (Properties of the Fréchet-derivative)**.**

i) The continuous linear operator L defining the Fréchet-derivative is unique.

ii) If F and G are Fréchet-differentiable in x_0, so are $F + G$ and λF for all

$\lambda \in \mathbb{R}$, *and it holds, that*

$$D_{x_0}(F+G) = D_{x_0}F + D_{x_0}G \qquad \text{and} \qquad D_{x_0}(\lambda F) = \lambda D_{x_0}F \ .$$

iii) If F and G are Fréchet-differentiable in x_0, then $F \cdot G$ is Fréchet-differentiable and the Fréchet-derivative satisfies the product rule:

$$D_{x_0}(F \cdot G) = G(x_0)D_{x_0}F + F(x_0)D_{x_0}G \ .$$

iv) Let X, Y, Z be normed spaces, and $F\colon D(F) \subset X \to Y$, as well as $G\colon D(G) \subset Y \to Z$ be Fréchet-differentiable operators with $F(D(F)) \subset D(G)$. Then $G \circ F$ is Fréchet-differentiable and the chain-rule applies:

$$D_{x_0}(G \circ F) = D_{F(x_0)}G \circ D_{x_0}F \ .$$

The first property allows to speak about "the" Fréchet-derivative. Also it allows to prove the remaining properties in the same way one does in the finite-dimensional case, by using the uniqueness of the linearization. If Fréchet-differentiability is given, the same rules apply to the Gâteaux-derivative and hence also for the first variation. Yet the opposite direction is not true.

By the similarities we have encountered so far, it is hardly a surprise that a vanishing Fréchet-derivative and thus vanishing variations for all h, is a necessary condition for extrema. Sufficient conditions, unlike in the finite-dimensional case, are however more complicated. In most cases, at least for physical theories, it is not important if a solution is minimizing or maximizing. Otherwise, usually there is an easy way to determine what kind of extremum it is, by physical reasoning.

G.2. Variation of Lagrange densities

In most physical applications the functionals of interest can be written as integration over a Lagrange density. A **Lagrange-density** $\mathcal{L}$ is an object, that maps functions to functions and thus defining itself a function of functions (and its variables):

$$\mathcal{L}(f(x), g(x), \ldots, x) = \ell(x) \ ,$$

where $\ell\colon U \subset \mathbb{R}^n \to \mathbb{R}$ is a proper function. Furthermore, the functionals $S(\phi)$ of interest often have Lagrange densities that do only depend on the components of the fields ${\phi^{a\ldots b}}_{c\ldots d}$ and their partial derivatives ${\phi^{a\ldots b}}_{c\ldots d,\mu}$:

$$\boxed{S(\phi) = \int_U \mathcal{L}({\phi^{a\ldots b}}_{c\ldots d}(x), {\phi^{a\ldots b}}_{c\ldots d,\mu}(x), x)\ d^n x.}$$

Here U has to be a compact subset of $\mathbb{R}^n$. In the case of invariant theories on Riemannian manifolds one can also use covariant derivative instead of partial derivatives. We shorten the notation of tensor components in the following, using $\phi^I_J \equiv {\phi^{a\ldots b}}_{c\ldots d}$.

Theorem G.2.1.
Let $S(\phi) = \int_U \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)\ d^n x$ be a Fréchet-differentiable functional with two times continuously differentiable Lagrangian. Then the first variation over fields, vanishing at the boundary ∂U, can be written as follows:

$$\delta S(\phi;\psi) = \int_U \left(\frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_J} - \sum_{\mu=1}^{n} \frac{\partial}{\partial \mu} \frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_{J,\mu}} \right) \psi^I_J\ d^n x\ .$$

Proof G.2.2.
Assuming Fréchet-differentiability we know that the Gâteaux-derivative exists and can simply calculate the first variation for the proof:

$$S(\phi + t\psi) - S(\phi) = \int_U \mathcal{L}(\phi^I_J + t\psi^I_J, \phi^I_{J,\mu} + t\psi^I_{J,\mu}, x) - \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)\ d^n x\ .$$

The difference of the Lagrange-densities can be Taylor-expanded for small t, which is given, when we take the limit.

$$\begin{aligned}\mathcal{L}(\phi^I_J + t\psi^I_J, \phi^I_{J,\mu} + t\psi^I_{J,\mu}, x) = \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x) \\ + t\ \frac{d}{dt}\bigg|_{t=0} \mathcal{L}(\phi^I_J + t\psi^I_J, \phi^I_{J,\mu} + t\psi^I_{J,\mu}, x) \\ + \mathcal{O}(t^2)\ .\end{aligned}$$

Thus the integrand becomes:

$$\begin{aligned}\mathcal{L}(\phi^I_J + t\psi^I_J, \phi^I_{J,\mu} + t\psi^I_{J,\mu}, x) - \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x) = t\ \frac{d}{dt}\bigg|_{t=0} \mathcal{L}(\phi^I_J + t\psi^I_J, \phi^I_{J,\mu} + t\psi^I_{J,\mu}, x) \\ + \mathcal{O}(t^2)\ .\end{aligned}$$

Evaluating the parameter derivative, using the parameter chain rule[1], we find

$$\frac{d}{dt}\bigg|_{t=0} \mathcal{L}(\phi^I_J + t\psi^I_J, \phi^I_{J,\mu} + t\psi^I_{J,\mu}, x) = \frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_J}\psi^I_J + \sum_{\mu=1}^{n} \frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_{J,\mu}}\psi^I_{J,\mu}\ .$$

Also we observe that

$$\frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_{J,\mu}}\psi^I_{J,\mu} = \frac{\partial}{\partial \mu}\left(\frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_{J,\mu}}\psi^I_J \right) - \left(\frac{\partial}{\partial \mu} \frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_{J,\mu}} \right) \psi^I_J\ .$$

Inserting our findings in the expression for $S(\phi + t\psi) - S(\phi)$ we get:

$$\begin{aligned}S(\phi + t\psi) - S(\phi) = t \int_U \left(\frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_J} - \sum_{\mu=1}^{n} \frac{\partial}{\partial \mu} \frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_{J,\mu}} \right) \psi^I_J\ d^n x \\ + t \int_U \frac{\partial}{\partial \mu}\left(\frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_{J,\mu}}\psi^I_J \right)\ d^n x + \mathcal{O}(t^2)\ .\end{aligned}$$

The integral in the second line can be rewritten using Stokes theorem (better known as Gauss-theorem or divergence theorem here). By assumption $\psi_J^I = 0$ on ∂U we find:

$$\int_U \frac{\partial}{\partial \mu}\left(\frac{\partial \mathcal{L}(\phi_J^I, \phi_{J,\mu}^I, x)}{\partial \phi_{J,\mu}^I}\psi_J^I\right)\, d^n x = \int_{\partial U} \frac{\partial \mathcal{L}(\phi_J^I, \phi_{J,\mu}^I, x)}{\partial \phi_{J,\mu}^I}\psi_J^I \, d\Omega = 0 \ .$$

Finally with $\delta S(\phi;\psi) = \lim_{t\to 0} \frac{1}{t} S(\phi + t\psi) - S(\phi)$ we get an expression for the first variation:

$$\delta S(\phi;\psi) = \int_U \left(\frac{\partial \mathcal{L}(\phi_J^I, \phi_{J,\mu}^I, x)}{\partial \phi_J^I} - \sum_{\mu=1}^{n} \frac{\partial}{\partial \mu}\frac{\partial \mathcal{L}(\phi_J^I, \phi_{J,\mu}^I, x)}{\partial \phi_{J,\mu}^I}\right)\psi_J^I \, d^n x \ .$$

□

It is common practice to imply the sum over μ by using ∂_μ instead of $\frac{\partial}{\partial \mu}$ in the sense of Ricci-calculus.

Lemma G.2.3 (Euler-Lagrange-equations)**.**
A smooth field ϕ extremizes the functional $S(\phi) = \int_U \mathcal{L}(\phi_J^I, \phi_{J,\mu}^I, x)\ d^n x$ with the boundary condition $\phi_J^I|_{\partial U} \equiv 0$ if it satisfies the ***Euler-Lagrange-equations*** *:*

$$\frac{\partial \mathcal{L}(\phi_J^I, \phi_{J,\mu}^I, x)}{\partial \phi_J^I} - \partial_\mu \frac{\partial \mathcal{L}(\phi_J^I, \phi_{J,\mu}^I, x)}{\partial \phi_{J,\mu}^I} = 0 \ .$$

Proof G.2.4.
All we need to do, is to show that $\delta S(\phi;\psi) = 0$, is equivalent to the Euler-Lagrange-equations. From the last theorem we know that

$$\delta S(\phi;\psi) = \int_U \left(\frac{\partial \mathcal{L}(\phi_J^I, \phi_{J,\mu}^I, x)}{\partial \phi_J^I} - \partial_\mu \frac{\partial \mathcal{L}(\phi_J^I, \phi_{J,\mu}^I, x)}{\partial \phi_{J,\mu}^I}\right)\psi_J^I \, d^n x \ .$$

If the Euler-Lagrange-equations are satisfied, we are integrating over the zero function, such that $\delta S(\phi;\psi) = 0$. The opposite direction is harder to show. Yet, there is an important theorem, called **Fundamental lemma of calculus of variations**, which states that an integrable function $f\colon \Omega \subset \mathbb{R}^n \to \mathbb{R}$ is identically zero, if

$$\int_\Omega f(x) g(x)\ d^n x = 0$$

for all $g \in C^\infty$ with compact support. Since we assumed U to be compact ψ_J^I has a compact support. □

[1] $\frac{d}{dt} f(x_1(t), \ldots, x_m(t)) = \sum_{i=1}^{m} \frac{\partial f(x_1(t), \ldots, x_m(t))}{\partial x_i} \frac{d}{dt} x_i(t).$

G.3. Physical conventions and notations

In the physical literature there are different notations and conventions for variational calculus. For example, arguments of functionals get square brackets $S[\phi]$. Most notably however, is the notation and concept of the functional derivative. As we have seen, the functional derivative (Fréchet-derivative) of a functional S defined by an integration over a Lagrange-density, can be written as integral operator

$$D_{\phi^I_J} S = \int_U d^n x \ \left(\frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_J} - \partial_\mu \frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_{J,\mu}} \right) .$$

The domain of the functional is the set of smooth tensor fields on a compact subset of $\mathbb{R}^n$. Also the Lagrange-density had to satisfy differentiability and dependence conditions. We notice, that the integral-operator defines a distribution over a function, commonly denoted by $\frac{\delta S}{\delta \phi^I_J}$:

$$\boxed{\frac{\delta S}{\delta \phi^I_J} = \frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_J} - \partial_\mu \frac{\partial \mathcal{L}(\phi^I_J, \phi^I_{J,\mu}, x)}{\partial \phi^I_{J,\mu}} .}$$

The term $\frac{\delta S}{\delta \phi^I_J}$ is also called **functional derivative**. This is possible, as a function defines a regular distribution by integration. However, we had to restrict the domain of the functional, that can be defined on a larger set. However, if one does so, the term $\frac{\delta S}{\delta \phi^I_J}$ defines no longer the Fréchet-derivative. We have seen so in the proof, where we could ignore boundary terms, which enter in a general setting. The physical jargon of this is: *the field variations vanish at the boundary.*

However, even with relaxed conditions, $\frac{\delta S}{\delta \phi^I_J}$ is called functional derivative in the physical literature. The reasoning behind this is, that although it does no longer define the Fréchet-derivative, it still defines the variation with respect to fields ψ, that have vanishing variations on the boundary:

$$\delta S(\phi, \psi) = \int_U \frac{\delta S}{\delta \phi^I_J} \psi^I_J \ d^n x .$$

The field ϕ is not necessary an extrmum anymore, since only variations in special directions vanish.

Also, the condition of U to be compact can be relaxed, as long as the variation fields ψ have a compact support.

Remark G.3.1.
More carefully one should be with concepts involving the delta function. An example for this would be the definition of functional derivatives:

$$\frac{\delta S}{\delta \phi^I_J(y)} = \lim_{t \to 0} \frac{1}{t} (S[\phi + t\delta(x-y)] - S[\phi]) .$$

As a means of notation, this works for Lagrange-densities. However, the delta-distribution is not regular. Hence, the above definition for functional derivative fails for more general functionals.

Index

Bibliography

[Alt12] A. Altland. *Advanced Quantum Mechanics*. Ed. by A. Altland. Vorlesungsskript. 2012. URL: http://www.thp.uni-koeln.de/alexal/pdf/advqm.pdf.

[AS10] A. Altland and B.D. Simons. *Condensed Matter Field Theory*. Cambridge books online. Cambridge University Press, 2010.

[Bau09] H. Baum. *Eichfeldtheorie, Eine Einführung in mathematische Konzepte*. Ed. by H. Baum. 2009. URL: https://www.mathematik.hu-berlin.de/~baum/Skript/KompaktkursHFB-05.pdf.

[BD94] J. Bjorken and S. Drell. *Relativistic Quantum Mechanics*. McGraw-Hill Education, 1994.

[Dun09] Gerald Dunne. *Functional Determinants in Quantum Field Theory*. Ed. by Gerald Dunne. 2009. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.218.814&rep=rep1&type=pdf.

[Fri03] K. Fritzsche. *Clifford-Algebren und Spin-Mannigfaltigkeiten*. Ed. by K. Fritzsche. 2003–2004. URL: http://www2.math.uni-wuppertal.de/~fritzsch/lect_mgf.html.

[HO07] S. Holtz and B. Obst. *Tensorprodukt*. Ed. by S. Holtz and B. Obst. Vorlesungsskript. 2007. URL: https://www-irm.mathematik.hu-berlin.de/~teschke/commalg/Tensor.pdf.

[Jän05] K. Jänich. *Vektoranalysis*. Springer-Lehrbuch. Springer, 2005.

[KW06] H. Kerner and W. von Wahl. *Mathematik für Physiker*. Springer Verlag, 2006.

[RW05] M. Roczen and H. Wolter. "Lineare Algebra individuell, Kapitel 4". In: *Lineare Algebra individuell*. 2005. URL: https://www.math.hu-berlin.de/~roczen/papers/tensor.pdf.

[RW08] A. Rosch and A. Weinkauf. *Quantenphysik*. Ed. by A. Rosch and A. Weinkauf. 2008.

[Sch09] D. Schliebner. *Lie-Gruppen und Lie-Algebren Eine Einführung*. Ed. by D. Schliebner. 2009. URL: https://www.mathematik.hu-berlin.de/~schliebn/dl/SoSe09-Lie%20Algebren.pdf.

[Sch08] F. Schwabl. *Quantenmechanik für Fortgeschrittene*. 5th ed. Springer Verlag, 2008.

[Sch13] F. Schwabl. *Quantenmechanik (QM I): Eine Einführung*. 7th ed. Springer-Lehrbuch. Springer Berlin Heidelberg, 2013.

[Swe15] G. Sweers. *Funktionentheorie*. Ed. by G. Sweers. Vorlesungsskript. 2015. URL: http://www.mi.uni-koeln.de/~gsweers/Skripte-in-PDF/Funktionentheorie2015.pdf.

[Wei14] Timo Weigand. *Quantum Field Theory I+II*. Ed. by Timo Weigand. 2014. URL: http://www.thphys.uni-heidelberg.de/~weigand/QFT2-14/SkriptQFT2.pdf.

[Wer11] D. Werner. *Funktionalanalysis*. Springer-Lehrbuch. Springer Berlin Heidelberg, 2011.

[Wik18] Wikipedia. *Complete set of commuting observables*. Ed. by Wikipedia. 2018. URL: https://en.wikipedia.org/wiki/Complete_set_of_commuting_observables.

[Zir10] M. Zirnbauer. *Lectures on Advanced Quantum Mechanics*. Ed. by M. Zirnbauer. 2010–2017. URL: http://www.thp.uni-koeln.de/zirn/011_Website_Martin_Zirnbauer/3_Teaching/LectureNotes/04AdvQM_WS10.pdf.

[Zir13] M. Zirnbauer. *Quantenphysik*. Ed. by M. Zirnbauer. 2013. URL: http://www.thp.uni-koeln.de/zirn/011_Website_Martin_Zirnbauer/3_Teaching/LectureNotes/All-QM_SS2013.pdf.

[Zir14] M. Zirnbauer. *Concepts of Quantum Field Theory*. Ed. by M. Zirnbauer. 2014. URL: http://www.thp.uni-koeln.de/zirn/011_Website_Martin_Zirnbauer/3_Teaching/LectureNotes/QFT1-SS2014.pdf.

[Zir15] M. Zirnbauer. *Klassische Theoretische Physik II*. Ed. by M. Zirnbauer. 2015–2016. URL: http://www.thp.uni-koeln.de/zirn/011_Website_Martin_Zirnbauer/3_Teaching/LectureNotes/KTP2_WS1516.pdf.

[Zir16] M. Zirnbauer. *Maxwell in Ketten*. Ed. by M. Zirnbauer. 2016. URL: http://www.thp.uni-koeln.de/zirn/011_Website_Martin_Zirnbauer/3_Teaching/LectureNotes/Maxwell-Chains_SS2016.pdf.