ANDREAS WEILER, TUM

RELATIVITY, PARTICLES, FIELDS
Beachten Sie bitte, dass das Skript noch nicht gründlich Korrektur gelesen wurde und es sicher noch Fehler gibt. Bitte schicken Sie sehr willkommene Vorschläge, Korrekturen, Kommentare an:

andreas.weiler@tum.de

Thanks to Tobias Duswald, Nepomuk Ritz, Patrick Selle, Ennio Salvioni, Stefan Schuldt, and Christian Schuster for helping to improve the script!

Der wackre Schwabe forcht' sich nit,
ing seines Weges Schritt vor Schritt.¹

¹ Ludwig Uhland oft zitiert von Albert Einstein.

Copyright © 2018 Andreas Weiler, TUM

Sommersemester 2017, Version vom December 31, 2018
Contents

1 Foreword 9
1.1 Units and conventions 9

2 Introduction 11
2.1 Why quantum field theory? 11
2.1.1 Necessity of the field viewpoint: causality 12
2.1.2 General features of relativistic quantum field theories 13
2.2 String Theory 13
2.2.1 The discrete chain 13
2.2.2 The continuous string 14
2.2.3 Quantum string theory 16
2.2.4 Review of the simple harmonic oscillator 17
2.2.5 Fock space of the discrete chain 18

3 Relativity 23
3.1 Galileo and Newton 23
3.1.1 Maxwell vs. Newton 24
3.1.2 Derivation of the Lorentz transformations 24
3.2 Consequences of special relativity 25
3.2.1 Relativity of simultaneity 26
3.2.2 Time dilation 27
3.2.3 Lorentz contraction 27
3.2.4 Addition of velocities 28
3.3 Minkowski space 29
3.3.1 Fourvectors 30
3.3.2 Minkowski Metric 31
3.3.3 Discrete transformations 32
3.3.4 Motion of a free particle 32
4 Classical Field Theory 37
4.1 Dynamics of fields 37
4.1.1 Example: the electro-magnetic field 37
4.1.2 The Lagrangian density 37
4.1.3 Example 1: the Klein-Gordon equation 39
4.1.4 Example 2: first order Lagrangians 39
4.1.5 Example 3: the Maxwell equations 40
4.1.6 Locality 40
4.1.7 Lorentz invariance 41
4.2 Symmetries: Noether’s Theorem and conserved currents 42
4.2.1 Conserved currents 43
4.2.2 Example 1: internal symmetries $U(1)$ 43
4.2.3 Example 2: an internal $SO(N)$ and a bit on Lie groups 44
4.2.4 Example 3: space-time translations 45
4.2.5 Summary of Noether’s theorem 47
4.3 Coulomb’s law and the propagator method 47
4.3.1 Fourier transformations 48
4.3.2 The Coulomb Potential 49
4.3.3 A non-linear example: the graviton Lagrangian 49
4.4 The Hamiltonian Formalism 51
4.4.1 Example: a real scalar field 52

5 Canonical Quantization 53
5.1 Free fields 53
5.1.1 The free scalar field 54
5.1.2 Infinites 55
5.1.3 Normal ordering 57
5.2 Particles 58
5.2.1 Multi-particle states and Bose statistics 59
5.3 Smeared operator valued distributions 60
5.3.1 Dimensional analysis and Coleman theorem 61
5.4 Some technicalities: relativistic normalization 62
5.4.1 Meaning of the state $\phi(x)|0\rangle$ 64
5.5 Complex scalar fields 64
5.6 Time dependent operators: the Heisenberg picture 66
5.7 Causal quantum fields 68
5.7.1 Propagators in quantum field theory 69
5.8 The Stückelberg-Feynman propagator 71
5.8.1 Green’s Functions 73
5.8.2 Particle creation by a classical source 74
5.9 Non-relativistic fields 76
5.9.1 A special case: quantum mechanics 78
5.9.2 Non-relativistic interactions 80

6 Interacting quantum fields 81
6.1 Dimensional analysis of interactions 81
6.1.1 Some examples of weakly coupled theories 82
6.2 The interaction picture 84
6.2.1 Dyson’s formula 85
6.3 A first look at scattering 87
6.3.1 Scattering simplified 88
6.3.2 Example: Meson Decay $\phi \to \bar{\psi}\psi$ 90
6.4 Wick’s Theorem 90
6.4.1 Example: nucleon scattering $\psi\psi \to \psi\psi$ 92
6.5 Diagrammatic Perturbation theory: Feynman Diagrams 94
6.5.1 Feynman rules 95
6.5.2 Revisiting $\psi\psi \to \psi\psi$ scattering 98
6.6 More scattering processes in the $\bar{\psi}\psi\phi$ theory 98
6.6.1 Nucleon anti-nucleon scattering: $\psi(p_1) + \bar{\psi}(p_2) \to \psi(q_1) + \bar{\psi}(q_2)$ 98
6.6.2 Meson-meson scattering $\phi\phi \to \phi\phi$ 99
6.6.3 Symmetry factors 99
6.7 Potentials 102
6.7.1 The Yukawa Potential 102
7.6 Perturbation Theory for Spinors 143
7.7 The Fermion Propagator 144
7.7.1 The Feynman propagator for Dirac fields 145
7.7.2 Nucleon-Nucleon scattering with spin 146
7.8 Feynman rules for Dirac Spinors 148
7.8.1 Example: Nucleon scattering $\Psi\Psi \rightarrow \Psi\Psi$ 149
7.8.2 Example: Nucleon-Anti-Nucleon scattering $\Psi\bar{\Psi} \rightarrow \Psi\bar{\Psi}$ 149
7.8.3 Example: Meson Scattering 149
7.8.4 Example: the Yukawa potential 150

8 Quantum Electro Dynamics 153
8.1 The Feynman Rules for Quantum Electrodynamics 153
8.2 Maxwell’s equations 154
8.3 Gauge Symmetry 155
8.3.1 Gauges 156
8.4 Quantizing the Electro-magnetic field 157
8.4.1 Coulomb gauge 157
8.4.2 Lorentz Gauge 159
8.4.3 The propagator 164
8.5 Coupling gauge fields to matter 164
8.5.1 Coupling photons to Dirac fermions 164
8.5.2 Coupling photons to charged scalars 165
8.6 The electric charge 166
8.7 Discrete symmetries of QED 166
8.7.1 Parity 167
8.7.2 Charge conjugation 167
8.7.3 Majorana neutrino 169
8.7.4 Furrys theorem 170
8.8 Time reversal 170
8.8.1 Quantum mechanics 170
8.8.2 Scalar Field 171
8.8.3 Spinor Field 172
8.9 Electron scattering and the Coulomb potential 172
8.9.1 Coulomb potential 173

9 Outlook 175
1

Foreword

These notes are not original but mostly a combination of results found in books and lecture notes. My contribution is here foremost in the selection and the perspective provided. I have made liberal use of the following references

- Peskin, Schroeder - *An Introduction to Quantum Field Theory* ★★★
- Schwartz - *Quantum Field Theory and the SM* ★★
- Mandll, Shaw - *Quantum Field Theory* ★
- Zee - *Quantum Field Theory in a Nutshell* ★
- Srednicki - *Quantum Field Theory* ★★★
- Weinberg - *Quantum Field Theory I* ★★★★
- Ramond - *Field Theory: a Modern Primer* ★★★
- Itzykson, Zuber - *Quantum Field Theory* ★★★

and many more. I strongly suggest that you find a book (or books) you like from the ones above and study it as a complement to these lecture notes.

This course provides a hopefully gentle introduction into the beautiful world of quantum field theory while it familiarizes you with relativity. It is the first real theory course in that it is the first course to tackle the theoretical framework that underlies all of nature. Let’s go.

1.1 Units and conventions

There are three fundamental dimensionful constants in nature: the speed of light $c$, Planck’s constant (divided by $2\pi$) $\hbar$, and Newton’s constant $G_N$. Their dimensions are

$$[c] = \text{length} \times \text{time}^{-1} \quad (1.1)$$
$$[\hbar] = \text{length}^2 \times \text{mass} \times \text{time}^{-1} \quad (1.2)$$
$$[G_N] = \text{length}^3 \times \text{mass}^{-1} \times \text{time}^{-2} \quad (1.3)$$
The whole point of units is that you can choose whatever units are most convenient! In particle physics and cosmology, we use "natural units"

\[ \hbar = c = 1 \]  

(1.4)

and so

\[ \text{length} = \text{time} = \text{mass}^{-1} = \text{energy}^{-1} \]  

(1.5)

we can express all dimensionful quantities in terms of a single scale which we choose to be mass or, equivalently, energy.\(^1\) I will temporarily reintroduce \(\hbar\) and \(c\) whenever convenient, mostly to show the classical \(\hbar \to 0\) limit or when we start with relativity. To convert the unit of energy back to units of length or time, we have to insert the relevant powers of \(c\) and \(\hbar\).

Energies will be given in units of eV (the electron volt)\(^2\) or more often GeV = \(10^9\) eV or TeV = \(10^{12}\) eV, since we are often dealing with high energies. Newton’s constant defines a mass scale

\[ G_N = M_P^{-2}, \quad M_P = 1.22 \times 10^{19} \text{GeV} \]  

(1.6)

this is the Planck scale. It corresponds to a length \(l_P \approx 10^{-33} \text{ cm}\).

Useful conversion factors are

\[ (1 \text{ GeV})^{-1}(\hbar c) = 0.1973 \text{ fm} \]  

(1.7)

\[ (1 \text{ GeV})/c^2 = 1.783 \times 10^{-24} \text{ g} \]  

(1.8)

\(\text{Table 1.1: Overview over the relevant fundamental scales of Nature.}\)
2

Introduction

2.1 Why quantum field theory?

It is the theory of the world. Here are the main arguments:

- All elementary particles look the same. An electron at the beginning of the universe has the exact same properties as an electron which is part of the atoms of your body. An excellent explanation for this is, if we assume that there is one field per elementary particle permeating the whole universe
  \[ \Psi^a = \Psi^a(x, t) \]  
  Elementary particles are then local, quantized excitations of this universal field. Each elementary particle corresponds to exactly one quantum field.

- Once we try to combine special relativity (SR) with quantum mechanics (QM), we find that particle number is not conserved anymore

  \[ QM + SR = QFT \]

  We will study Dirac’s argument for anti-particles in detail later. Let it just suffice to say, that if we try to localize a particle in a box smaller than
  \[ \lambda = \frac{h}{mc} \]  
  then the energy uncertainty is of order of the mass of the particle. To see this, start with the Heisenberg relation for the momentum uncertainty of a particle in a box of size \( L \)
  \[ \Delta p \geq \frac{h}{L} \]

  Relativistically, momentum and energy are on the same footing
  \[ \Delta E \geq \frac{hc}{L}. \]

  But if the uncertainty exceeds \( \Delta E = 2mc^2 \) then we can create particle anti-particle pairs out of the vacuum.

  The Compton wavelength is always smaller than the de Broglie wavelength \( \lambda_{dB} = \frac{h}{p} \). If you like, the de Broglie wavelength is the distance at which the wavelike nature of particles becomes apparent; the Compton wavelength is the distance at which the concept of a single point like particle breaks down completely.

This is the so-called Compton wavelength.
2.1.1 Necessity of the field viewpoint: causality

We will show now that causality makes a multi-particle theory necessary. We start with the amplitude for a particle propagating from $x_0$ to $x$ during the time $t$

$$U(t) = \langle x | e^{-iHt} | x_0 \rangle$$  \hspace{1cm} (2.3)

The energy of a free particle in non-relativistic quantum mechanics is $E = p^2/(2m)$ so

$$U(t) = \langle x | e^{-i\vec{p}^2/(2m)t} | x_0 \rangle$$  \hspace{1cm} (2.4)

$$= \int \frac{d^3p}{(2\pi)^3} \langle x | e^{-i\vec{p}^2/(2m)t} | p \rangle \langle p | x_0 \rangle$$  \hspace{1cm} (2.5)

$$= \frac{1}{(2\pi)^3} \int d^3p e^{-i\vec{p}^2/(2m)t} \langle x | p \rangle \langle p | x_0 \rangle$$  \hspace{1cm} (2.6)

$$= \frac{1}{(2\pi)^3} \int d^3p e^{-i\vec{p}^2/(2m)t} e^{i\vec{p} \cdot (x-x_0)}$$  \hspace{1cm} (2.7)

$$= \left( \frac{m}{2\pi it} \right)^{3/2} \exp \left( \frac{im(x-x_0)^2}{2t} \right)$$  \hspace{1cm} (2.8)

where in the last step we have used a general gaussian integral with $A = it/m \mathbb{1}$ and $B = i(x - x_0)$. This amplitude is non-zero for all $x$ and $t$, which implies that a particle can propagate between any two points in an arbitrarily short time. In a relativistic theory, no signal can propagate faster than light and this would mean that causality is violated! We can try to amend the above formula by using the relativistic Pythagoras for the energy

$$E^2 = (pc)^2 + (mc^2)^2$$  \hspace{1cm} (2.9)

and taking the positive root

$$U(t) = \langle x | e^{-i\sqrt{p^2+m^2}t} | x_0 \rangle$$

$$= \int \frac{d^3p}{(2\pi)^3} \langle x | e^{-i\sqrt{p^2+m^2}t} | p \rangle \langle p | x_0 \rangle$$

$$= \frac{1}{(2\pi)^3} \int d^3p e^{-i\sqrt{p^2+m^2}t} e^{i\vec{p} \cdot (x-x_0)}$$

$$= \frac{1}{(2\pi)^3} \int_0^\infty dp 2p \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta \ e^{-i\sqrt{p^2+m^2}t} e^{ip|x-x_0| \cos \theta}$$

$$= \frac{1}{(2\pi)^3} \int_0^\infty dp 2p e^{-i\sqrt{p^2+m^2}t} \int_{-1}^1 d \cos \theta e^{ip|x-x_0| \cos \theta}$$

$$= \frac{1}{2\pi^2 \ |x-x_0|} \int_0^\infty dp e^{-i\sqrt{p^2+m^2}t} \sin(p|x-x_0|)$$

where we have chosen the z-axis along the $x - x_0$ direction. We now want an approximate solution for this integral for $x^2 \gg t^2$ (far outside the light-cone). We will use the stationary phase approximation. The phase function

$$\phi(p) \approx px - t \sqrt{p^2 + m^2}$$  \hspace{1cm} (2.10)

where $x - x_0$ is far outside the light-cone.

2. Inserting the identity

$$1 = \int \frac{d^3p}{(2\pi)^3} \langle p | p \rangle$$

and

$$\langle x | p \rangle = e^{ip \cdot x}$$

3 A very useful formula:

$$\int d^n x e^{-\frac{1}{2} \sum_{i,j=1}^n \Lambda_{ij} x_i x_j + \sum_{i=1}^n B_i x_i}$$

$$= \int d^n x e^{-\frac{1}{2} x^T A x + B^T x}$$

$$= \sqrt{\frac{(2\pi)^n}{\det A}} \exp \left( \frac{1}{2} B^T A^{-1} B \right)$$

4 More on that later.

5 Why?

The main idea of stationary phase methods relies on the cancellation of sinusoids with rapidly varying phase. If many sinusoids have the same phase and they are added together, they will add constructively. If, however, these same sinusoids have phases which change rapidly as the frequency changes, they will add incoherently, varying between constructive and destructive addition at different times.

has a stationary point at
\[
\frac{\partial^2 \phi}{\partial p^2} \bigg|_{p_0} = 0 \quad \Rightarrow \quad p_0 = \frac{i mx}{\sqrt{x^2 - t^2}} \quad (2.11)
\]
Substituting this value back for \( p \), we find that up to a rational function of \( x \) and \( t \), that
\[
U(t) \sim e^{i\phi(p_0)} \sim e^{-m\sqrt{x^2 - t^2}} \quad (2.12)
\]
We find a small, exponentially suppressed amplitude but still non-zero outside the light-cone, **causality is still violated**.

We see from the exponential suppression that causality is violated if we get to distances as small as \( x \sim 1/m \), for distances \( x \gg 1/m \) there is a negligible chance to find the particle outside the light-cone\(^6\), so at distances much greater than the Compton wavelength of a particle, the single-particle theory will not lead to measurable violations of causality. This is in accordance with our earlier arguments based on the uncertainty principle: multi-particle effects become important when you are working at distance scales of order the Compton wavelength of a particle.

Quantum field theory finds a miraculous solution to the causality problem: the propagation across a space-like interval is indistinguishable from the propagation of an anti-particle in the opposite direction. We will find that the amplitudes for particle and anti-particle propagation exactly cancel outside the light-cone – causality is preserved!

2.1.2 General features of relativistic quantum field theories

- CPT
- Spin-statistics
- Very constraining for higher spins, only \( S \leq 2 \) theories.
- Massless particles \((m = 0)\): \( S = 1 \) gauge theory (like electromagnetism), \( S = 2 \) gravity For the massless spin-2 field, the requisite gauge principle can be shown to be general covariance, which leads to Einstein’s theory, if you want to couple \( S = 3/2 \) to gravity: unique theory super-gravity

2.2 String Theory

2.2.1 The discrete chain

We end up with a quantum field by quantizing a classical field. A simple example is the classical string. We will define our string as a long-wavelength limit (= low energy limit) of a harmonically coupled discrete chain. We take \( N + 2 \) masses with a quadratic nearest-neighbor coupling
\[
L(q, \dot{q}) = T - V = \sum_{j=0}^{N+1} \left[ \frac{m}{2} \dot{q}_j^2 - \frac{K}{2} (q_j - q_{j+1})^2 \right] \quad (2.13)
\]
You can also solve this integral exactly in terms of Bessel-functions and convince yourself of the veracity of this statement.

\(^6\) Light-cone: \( x^2 = c^2 t^2 \), with temporarily reintroduced \( c \).
The coordinate \( q_j(t) \) is the longitudinal displacement of the \( j \)-th mass along the one-dimensional chain. The equilibrium distance \( a \) between masses sets the rest location of the masses

\[
x_j \equiv j \cdot a \tag{2.14}
\]

At the endpoints, the chain is fixed:

\[
q_0(t) = q_{N+1}(t) \equiv 0 \tag{2.15}
\]

The remaining \( N \) masses satisfy the equation of motion (EOM)

\[
m \ddot{q}_j - \kappa (q_{j+1} - 2q_j + q_{j-1}) = 0 \tag{2.16}
\]

for \( j = 1, \ldots, N \). The normal modes diagonalize the EOMs and have the form

\[
q_j(t) = \cos(\omega t) \sin(jp) \tag{2.17}
\]

The boundary-conditions in Eq. (2.15) require \( p \) to be

\[
p_n = \frac{\pi n}{N + 1} \tag{2.18}
\]

for \( n = 1, \ldots, N \). Substituting this into the EOMs, we get the dispersion relations for the \( N \) independent normal frequencies \( \omega_n \):

\[
\omega_n = \omega_0 \sin \left( \frac{\pi n}{2(N + 1)} \right), \quad \omega_0 = 2 \sqrt{\frac{\kappa}{m}} \tag{2.19}
\]

The frequency \( \omega_0 \) is the cut-off frequency since the modes with \( n > N \) are merely repeating the lower ones.

### 2.2.2 The continuous string

The total length of the chain is

\[
R = Na \tag{2.20}
\]

We find the continuum limit by taking the mass distance \( a \) to zero while keeping the the total length \( R \) fixed

\[
a \to 0, \quad N \to \infty, \quad R = Na = \text{const.} \tag{2.21}
\]

In this limit the discrete chain approaches a continuous string and the labeled coordinates \( q_j(t) \) approach a classical field defined as

\[
q(x, t) \equiv q_j(t) \tag{2.22}
\]

We can take the limit also for the Lagrangian by replacing

\[
(q_j - q_{j+1})^2 \to \left( \frac{q_j - q_{j+1}}{a} \right)^2 \to a^2 \left( \frac{\partial q(x, t)}{\partial x} \right)^2 \tag{2.23}
\]

\[
\sum_j \frac{1}{a} \sum_j a \to \frac{1}{a} \int_0^R dx \tag{2.24}
\]

Keeping the mass density \( \rho \) and string tension \( \sigma \) finite, we define

\[
\rho = \frac{m}{a}, \quad \sigma = \kappa a \tag{2.25}
\]
to obtain the Lagrangian

$$L = \frac{1}{2} \int_0^R dx \left[ \rho \left( \frac{\partial q(x, t)}{\partial t} \right)^2 - \sigma \left( \frac{\partial q(x, t)}{\partial x} \right)^2 \right]$$

(2.26)

and the EOM

$$\frac{1}{c^2} \frac{\partial^2 q(x, t)}{\partial t^2} - \frac{\partial^2 q(x, t)}{\partial x^2} = 0$$

(2.27)

which is of course a wave-equation with a speed of sound of

$$c^2 = \frac{\sigma}{\rho}$$

(2.28)

General solutions can be expanded in

$$q(x, t) = a \cdot e^{i(kx - \omega t)} + b \cdot e^{-i(kx - \omega t)}$$

(2.29)

with a dispersion law

$$\omega = c k$$

(2.30)

As above, the boundary conditions of fixed ends

$$q(0, t) = q(R, t) = 0$$

(2.31)

give us a form of the normal modes

$$q_n(x, t) = \cos(\omega_n t) \sin(k_n x)$$

(2.32)

with $$\omega_n = c k_n$$ and

$$k_n = \frac{\pi n}{R}$$

(2.33)

The normal-modes reproduce the discrete chain as long as $$n/N \ll$$

![Figure 2.2: Normal modes of a classical string $$\omega_n \sim n$$ and a discrete chain for $$N=8$$ which has $$\omega_n \sim \sin(n/N)$$.](image)

1. We find however that the number of modes of the continuous string is infinite! Only the first $$N$$ modes correspond to those of the discrete chain. Therefore, we see that there is a **cut-off frequency** $$\omega_C$$ which is

$$\omega_C \equiv \omega_N = \frac{\pi c}{a} \sim 1/a$$

(2.34)
This is of the same parametric size as the maximum frequency of the
discrete chain \( \omega_0 = 2\sqrt{\frac{c}{m}} = 2 \frac{c}{a} \), see Eq. (2.19). Our continuum
model \( q(x,t) \) is a good representation of the discrete system only for
\( \omega \ll \omega_C \).

This is a first encounter with the important concept of \textbf{renormalization}. The cut-off frequency
\[ \omega_C \sim \frac{1}{a} \]
is a theoretical necessity. The cut-off tells us until which scale, we can trust our continuum theory. A continuous string at frequencies \( \omega \gg \omega_C \) is not a good description of the discrete chain: the minimal wavelength is set by the
atomic distance, below which there is nothing to oscillate. Without a cut-off the specific heat of the string would
diverge since at a temperature \( T \) each mode (equipartition theorem in statistical physics) contributes an amount
\( kT \), where \( k \) is the Boltzmann constant.

Here, we cannot determine the value of the cut-off from the long-wavelength theory alone, because only the
combination
\[ c \sim a \omega_C \]
appears. We can absorb the cut-off in measurable parameters, as in Eq. (2.25). This is called \textit{renormalization}. If
this can be done, a theory is called \textit{renormalizable}. You will discuss this at length in the chapter of loop correc-
tions in quantum field theories.

If a theory is \textit{non-renormalizable}, the behavior can be sensitive to short-distance physics, which in this case would
be the atomic motion. This behavior does appear random on macroscopic scales, as in the propagation of a crack
or the nucleation of a raindrop.

2.2.3 Quantum string theory

We quantize the classical continuum theory to obtain a first example
of a quantum field theory. The Hamiltonian of the discrete chain is
\[ H(p,q) = T + V = N \sum_{j=1}^{N} \left[ \frac{p_j^2}{2m} + \frac{\kappa}{2} (q_j - q_{j+1})^2 \right] \] (2.35)
with \( p_j = m \dot{q}_j \). We can quantize the system, if we can replace \( p_j \) and \( q_j \) with hermitian operators which satisfy the commutation
\[ \left[ \hat{p}_j, \hat{q}_k \right] = -i\hbar \delta_{jk} \] (2.36)

We will switch to periodic boundary conditions now,
\[ \hat{q}_{j+N}(t) = \hat{q}_j(t) \] (2.37)
since they allow us to use just the exponentials
\begin{align*}
\hat{q}_j(t) &= \frac{1}{\sqrt{N}} \sum_{n=-N/2}^{N/2} \hat{Q}_n e^{i2\pi nj/N} \\
\hat{p}_j(t) &= \frac{1}{\sqrt{N}} \sum_{n=-N/2}^{N/2} \hat{P}_n e^{i2\pi nj/N}
\end{align*} (2.38, 2.39)
where the operators \( \hat{P}_n \) and \( \hat{Q}_n \) satisfy
\begin{align*}
\left[ \hat{P}_n, \hat{Q}_m \right] &= -i \delta_{nm}, \\
\hat{P}_n &= \hat{P}_{-n} \\
\hat{Q}_n &= \hat{Q}_{-n}
\end{align*} (2.40, 2.41, 2.42)

In the following we will mostly set \( \hbar = 1 \).

Think of the system as being on a
torus. The spacing of the frequencies
is doubled but each frequency ap-
ppears two-fold. We are not interested
in the behaviour at the endpoints
and we will be at large \( N \).

We will switch to periodic boundary conditions now,
The Hamiltonian is now simplified to a sum of independent harmonic oscillators

$$H = \sum_{n=1}^{N} \left[ \frac{1}{2m} \hat{P}_n^2 \hat{P}_n + \frac{1}{2} \hbar \omega_n^2 \hat{Q}_n^2 \right]$$

(2.43)

with eigen-frequencies

$$\omega_n^2 = \frac{4k}{m} \sin^2 \left( \frac{\pi n}{N} \right)$$

(2.44)

As you remember from your QM course, we can write the energy-eigenvalues of the harmonic oscillator with frequency $\omega$ as

$$\epsilon_n = \hbar \omega \left( \frac{1}{2} + n \right), \quad n = 0, 1, \ldots$$

(2.45)

where $n$ is the **occupation number**. We can therefore write the energy-eigenvalues of $H$ as a set of occupation numbers

$$E_\alpha = \frac{N}{2} \sum_{n=-N/2}^{N/2} \omega_n \left( \frac{1}{2} + \alpha_n \right)$$

(2.46)

with $\alpha_n = 0, 1, 2, \ldots$

### 2.2.4 Review of the simple harmonic oscillator

Consider the quantum mechanical Hamiltonian

$$H = \frac{1}{2} \hat{p}^2 + \frac{1}{2} \hbar \omega^2 \hat{q}^2$$

(2.47)

with the canonical commutation relations

$$[\hat{q}, \hat{p}] = i$$

(2.48)

To find the spectrum we define the creation and annihilation operators (also known as raising/lowering operators, or sometimes ladder operators)

$$a = \sqrt{\frac{\omega}{2}} q + \frac{i}{\sqrt{2\omega}} p,$$

$$a^\dagger = \sqrt{\frac{\omega}{2}} q - \frac{i}{\sqrt{2\omega}} p$$

(2.49, 2.50)

which we can invert to

$$q = \frac{1}{\sqrt{2\omega}} (a + a^\dagger),$$

(2.51)

$$p = -i \sqrt{\frac{\omega}{2}} (a - a^\dagger)$$

(2.52)

Substituting the expressions we find

$$[a, a^\dagger] = 1$$

(2.53)
while the Hamiltonian is
\[ H = \frac{1}{2} \omega (aa^\dagger + a^\dagger a) \] (2.54)
\[ H = \omega (a^\dagger a + \frac{1}{2}) \] (2.55)

One can easily confirm that the commutators between the Hamiltonian and the creation and annihilation operators are given by
\[ [H, a^\dagger] = \omega a^\dagger, \] (2.56)
\[ [H, a] = -\omega a \] (2.57)

Let \(|E\rangle\) be an eigenstate with energy \(E\), so that
\[ H|E\rangle = E|E\rangle \] (2.58)

Then we can construct more eigenstates by operating with \(a\) and \(a^\dagger\),
\[ Ha^\dagger|E\rangle = (E + \omega)a^\dagger|E\rangle, \] (2.59)
\[ Ha|E\rangle = (E - \omega)a|E\rangle \] (2.60)

So we find that we can generate a ladder of states for each \(k\) with energies
\[ \ldots, E - \omega, E, E + \omega, E + 2\omega, \ldots \] (2.61)

The energy is bounded from below and there must be a ground state \(|0\rangle\) which satisfies
\[ a|0\rangle = 0 \] (2.62)

This has the ground-state energy (also called zero-point energy),
\[ H|0\rangle = \frac{1}{2}\omega|0\rangle \] (2.63)

2.2.5 Fock space of the discrete chain

For this simple system, the Hilbert space can be identified with the tensor product of the Hilbert spaces of the individual atoms or masses. Thus, if we identify with \(|\Psi\rangle_n\) an arbitrary state in the Hilbert space of the \(n\)-th atom, the states of the chain can then be written as
\[ |\Psi\rangle = |\Psi\rangle_1 \otimes |\Psi\rangle_2 \otimes \ldots \otimes |\Psi\rangle_N \] (2.64)

This is a first example of the Fock space of multi-particle systems.

We can further diagonalize the Hamiltonian by making use of the creation and annihilation operator formalism for the harmonic oscillator.
\[ a_k \equiv \sqrt{\frac{m\omega_k}{2}} (Q_k + \frac{i}{m\omega_k} P_k) \] (2.65)
\[ a_k^\dagger \equiv \sqrt{\frac{m\omega_k}{2}} (Q_{-k} - \frac{i}{m\omega_k} P_{-k}) \] (2.66)
With this definition we find that the ladder operators obey the
commutation relations.

\[
[a_k, a_{k'}^\dagger] = \delta_{kk'} \quad (2.67)
\]
\[
[a_k, a_{k'}] = [a_k^\dagger, a_{k'}^\dagger] = 0 \quad (2.68)
\]

**Proof:** Plugging the definitions of \(a_k\) and \(a_k^\dagger\) into the commutator
and using the fact that the commutator is bilinear, we can calculate\(^7\) Thanks to Nepomuk Ritz!

\[
[a_k, a_{k'}^\dagger] = \left[ \sqrt{\frac{\omega_k \omega_{k'}}{2}} \left( Q_k + \frac{i}{m \omega_k} P_k \right), \sqrt{\frac{\omega_{k'} \omega_{k'}}{2}} \left( Q_{k'} - \frac{i}{m \omega_{k'}} P_{k'} \right) \right]
\]
\[
= \frac{m \sqrt{\omega_k \omega_{k'}}}{2} \left( [Q_k, Q_{k'}] - \frac{i}{m \omega_{k'}} [Q_k, P_{k'}] + \frac{i}{m \omega_{k}} [P_k, Q_{k'}] + \frac{1}{m^2 \omega_k \omega_{k'}} [P_k, P_{k'}] \right)
\]
\[
= \frac{m \sqrt{\omega_k \omega_{k'}}}{2} \left( \frac{i}{m \omega_{k'}} [P_{k'}^\dagger, Q_k] - \frac{i}{m \omega_{k}} \left( [P_{k'}^\dagger, Q_{k'}] \right)^\dagger \right)
\]
\[
= \frac{m \sqrt{\omega_k \omega_{k'}}}{2} \left( \delta_{kk'} \frac{2}{m \omega_{k}} \right) = \frac{m \omega_k}{2} \delta_{kk'} = \delta_{kk'}
\]

where we have also used that \(P_{-k} = P_{k}^\dagger, Q_{-k} = Q_{k}^\dagger\) and that \([P_{k'}^\dagger, Q_{k}] = -i \delta_{kk'}\).

\[
[a_k, a_{k'}] = \left[ \sqrt{\frac{\omega_k \omega_{k'}}{2}} \left( Q_k + \frac{i}{m \omega_k} P_k \right), \sqrt{\frac{\omega_{k'} \omega_{k'}}{2}} \left( Q_{k'} + \frac{i}{m \omega_{k'}} P_{k'} \right) \right]
\]
\[
= \frac{m \sqrt{\omega_k \omega_{k'}}}{2} \left( [Q_k, Q_{k'}] + \frac{i}{m \omega_{k'}} [P_k, Q_{k'}] + \frac{i}{m \omega_{k}} [Q_k, P_{k'}] - \frac{1}{m^2 \omega_k \omega_{k'}} [P_k, P_{k'}] \right)
\]
\[
= \frac{m \sqrt{\omega_k \omega_{k'}}}{2} \left( \delta_{kk'} \frac{2}{m \omega_{k'}} \right) = \delta_{kk'}.
\]

Since the commutator is antisymmetric, we have \([Q_k, P_{k'}] = -[P_{k'}, Q_k]\) and therefore, upon exchanging the indices \(k\) and \(k'\)
for this term (which we can do, because \([P_{k'}, Q_k] = -i \delta_{k, k'}\)), the
remaining terms cancel and we have

\[
[a_k, a_{k'}] = 0.
\]

If we take the hermitian conjugate of this equation, we immediately get

\[
[a_k, a_{k'}]^\dagger = [a_{k'}^\dagger, a_k^\dagger] = [a_{k'}^\dagger, a_k^\dagger] = 0
\]
and therefore also

\[
[a_k^\dagger, a_{k'}^\dagger] = 0.
\]

which completes the proof.

The Hamiltonian becomes

\[
H = \sum_{k=-N/2}^{N/2} \omega_k \left( a_k^\dagger a_k + \frac{1}{2} \right) \quad (2.69)
\]
Further we can show that
\[ [H, a_k] = -\omega_k a_k \]  
\[ [H, a_k^\dagger] = \omega_k a_k^\dagger \]
and therefore these operators take us between energy eigenstates.

Let \(|E\rangle\) be an eigenstate with energy \(E\), so that
\[ H|E\rangle = E|E\rangle \]
Then we can construct more eigenstates by operating with \(a_k\) and \(a_k^\dagger\),
\[ H a_k^\dagger |E\rangle = \sum_q \omega_q (a_q^\dagger a_q + 1/2) a_k^\dagger |E\rangle \]
So we find that we can generate a ladder of states for each \(k\) with energies
\[ \ldots, E - \omega_k, E + \omega_k, E + 2\omega_k, \ldots \]
The energy is bounded from below and there must be a ground state \(|0\rangle\) which satisfies
\[ a_k|0\rangle = 0 \quad \text{for all } k \]
This has the ground-state energy (also called zero-point energy),
\[ H|0\rangle = \frac{1}{2} \sum_k \omega_k |0\rangle \]
Excited states then can be generated by repeated application of \(a_k\),
\[ |n_1, n_2, \ldots, n_N\rangle = (a_1^\dagger)^{n_1}(a_2^\dagger)^{n_2} \cdots (a_N^\dagger)^{n_N}|0\rangle \]
where we have not been careful about normalization, so \(|n|n\rangle \neq 1\).

We now again move to the continuum limit with \(x = j \cdot a\) to find
\[ H_{\text{cont}} = \int_0^R dx \left[ \frac{1}{2\rho} \tilde{\pi}^2(x, t) + \frac{\sigma}{2} \left( \frac{\partial \hat{\phi}(x, t)}{\partial x} \right)^2 \right] \]
and the conjugate momentum operator
\[ \hat{\pi}(x, t) = \frac{\hat{p}_j(t)}{a} = \rho \frac{\partial \hat{j}_j(t)}{\partial t} = \rho \frac{\partial \hat{\phi}(x, t)}{\partial t} \]
The quantum field \(\hat{\phi}(x, t)\) and its conjugate momentum \(\hat{\pi}(x, t)\) satisfy the equal-time commutation relation
\[ [\hat{\phi}(x, t), \hat{\pi}(x', t)] = i\hbar \delta(x - x') \]
where I have temporarily reintroduced \(\hbar\) to show how to recover the classical limit \(\hbar \to 0\). Also for quantum string we have to introduce a cut-off frequency \(\omega_C\).
We have seen that in particular for distances smaller than $x \sim 1/m$ or energies above $E \sim m$, we need to extend the QM framework because of causality violation and the possibility of particle creation. These energies are however also energies where relativistic effects become important. Therefore we will spend some time on the basics of relativity in the next chapter.

We started with a discrete chain for pedagogical reasons. Of course we do not believe that the fields observed in Nature, such as the electron field or the photon field, are actually constructed of point masses tied together with springs. A modern picture which was first used by Landau-Ginzburg and fully understood by S. Weinberg, is that we start with the desired symmetry, say Lorentz invariance or a conserved $U(1)$ charge (more on both later) decide on the fields we want by specifying how they transform under the symmetry (in this case we decided on a scalar field $\phi(x,t)$ as the local displacement) and then write down the action involving no more than two time derivatives – because we don’t know how to quantize actions with more than two time derivatives (we usually end up with ghosts). See also the discussion in the statistical physics script (in particular the chapter: Landau-Ginzburg Theorie)
3

Relativity

3.1 Galileo and Newton

Two observers, one on a moving train \((x, y, z)\) with velocity \(u\) and one on the ground \((x', y', z')\), can be related by

\[
\begin{align*}
t' &= t \\
x' &= x + u \cdot t \\
y' &= y \\
z' &= z
\end{align*}
\] (3.1) (3.2) (3.3) (3.4)

which are known as the Galilean transformations. The differential form is

\[
\begin{align*}
dt' &= dt \\
dx' &= dx + u \cdot dt
\end{align*}
\] (3.5) (3.6)

We can add velocities trivially: e.g. a ball with \(3 \text{ m/s}\) on a train with \(20 \text{ m/s}\) will be observed on the ground as \(3 \text{ m/s} + 20 \text{ m/s} = 23 \text{ m/s}\). We see this also with the Galilean transformations

\[
v' = \frac{dx'}{dt'} = \frac{dx}{dt} + u = v + u
\] (3.7)

We now see that the invariance of Newtonian mechanics under Galilean transformations follows merely because Newton’s law involves 2nd derivatives, so

\[
m \frac{d^2 x'}{dt'^2} = m \frac{dx}{dt} + u = m \frac{d^2 x}{dt^2}
\] (3.8)

We also see importantly where Galilean invariance in Newtonian mechanics fails: if \(u\) changes in magnitude or direction with time then we find\(^1\)

\[
F' = ma' = ma + m \frac{du}{dt}
\] (3.9)

we feel this as an additional force, e.g. as a centrifugal force in an accelerated frame of reference. Let us check that the Newtonian action is invariant, too

\[
S = \int dt' \frac{1}{2} m \left( \frac{dx'}{dt'} \right)^2 = \int dt \frac{1}{2} m \left( \frac{dx}{dt} + u \right)^2
\] (3.10)

\[
= \int dt \frac{1}{2} m \left( \frac{dx}{dt} \right)^2 + u \int dt m \frac{dx}{dt} + u^2 \int dt \frac{1}{2} m
\] (3.11)

\(^1\) All inertial reference frame is a frame in which Newton's first axiom holds, i.e. every object on which no force is acting either stays at rest or moves uniformly.
for fixed initial and final conditions, the second term is just an
irrelevant additional constant, like the last term. We can generalize
this to a many particle system with interactions

\[ S = \int dt \left\{ \sum_{a} \frac{1}{2} m_a \left( \frac{dx_a}{dt} \right)^2 - \sum_{a \neq b} V(x_a - x_b) \right\} \] (3.12)

Note, that it is necessary for the interaction potential to just de-
pend on the difference \( x_a - x_b \). We also see that all observers have
to agree on the same mass \( m \), there is no \( m' \) otherwise Galilean
transformations would not work.

### 3.1.1 Maxwell vs. Newton

Maxwell’s equations are not invariant under Galilean transforma-
tions. In particular, the velocity of light is the same in all inertial
systems

\[ c' = c \] (3.13)

contradicting Eq. (3.7). Various eminent physicists in the late 19th
century tried to reconcile this fact by postulating that light – like
sound – had to propagate in a medium, the aether. In 1887 (when
Einstein was eight years old) Michelson and Morley performed a
famous experiment to detect the aether and failed.

#### The Michelson–Morley experiment

The Michelson–Morley experiment was performed over the spring and summer of 1887 by Albert A. Michelson and Edward W. Morley at what is now Case Western Reserve University in Cleveland, Ohio, and published in November of the same year. It compared the speed of light in perpendicular directions, in an attempt to detect the relative motion of matter through the stationary luminiferous aether ("aether wind"). The result was negative, in that the expected difference between the speed of light in the direction of movement through the presumed aether, and the speed at right angles, was found not to exist; this result is generally considered to be the first strong evidence against the then-prevalent aether theory, and initiated a line of research that eventually led to special relativity, which rules out a stationary aether.

Michelson–Morley type experiments have been repeated many times with steadily increasing sensitivity. These include experiments from 1902 to 1905, and a series of experiments in the 1920s. More recent optical resonator experiments confirmed the absence of any aether wind at the \( 10^{-15} \) level. Together with the Ives–Stilwell and Kennedy–Thorndike experiments, Michelson–Morley type experiments form one of the fundamental tests of special relativity theory. https://en.wikipedia.org/wiki/Michelson-Morley_experiment

### 3.1.2 Derivation of the Lorentz transformations

We consider two inertial systems and assume that they coincide for
\( t = 0 = t' \) with a relative velocity \( v \).

Let us know see how we can modify the Galilean transformations
so that a spherical light wave emanating from the origin

\[ (ct)^2 = x^2 + y^2 + z^2 \] (3.14)

does not depend on the observer and is the same in the moving
inertial system

\[ (ct')^2 = x'^2 + y'^2 + z'^2 \] (3.15)
Let us assume that the origin is moving along the $x$-axis. We know then
\[ y = y', \quad z = z' \]  
(3.16)
The most general linear solution is, as can be verified by direct substitution using (H1)
\[
\begin{pmatrix} x \\ ct \end{pmatrix} = \begin{pmatrix} \cosh \eta & \sinh \eta \\ \sinh \eta & \cosh \eta \end{pmatrix} \begin{pmatrix} x' \\ ct' \end{pmatrix} 
\]  
(3.17)
This should remind you of an Euclidean rotation which leaves $x^2 + y^2$ invariant (instead of $(ct)^2 - x^2$).

To find the role of $\eta$ in the physical setting, record the origin’s progression, i.e. $x' = 0, x = vt$. The equations become (using first $x' = 0$),
\[ x = ct' \sinh \eta, \]
\[ ct = ct' \cosh \eta. \]  
(3.18) (3.19)

Now divide :
\[ \frac{x}{ct} = \tanh \eta = \frac{v}{c} = \beta \Rightarrow \]
\[ \sinh \eta = \frac{v}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad \cosh \eta = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, \]
where $x = vt$ was used in the first step, (H2) and (H3) in the second, which, when plugged back in (3.17), gives
\[ x = \frac{x' + vt'}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad t = \frac{t' + \frac{v}{c^2} x'}{\sqrt{1 - \frac{v^2}{c^2}}}, \]  
(3.20)
or, with the usual abbreviations,
\[
\begin{align*}
x &= \gamma (x' + vt'), \\
t &= \gamma \left( t' + \frac{vx'}{c^2} \right),
\end{align*}
\]  
(3.21) (3.22)

We see that the transformations for time and space are now symmetric!\(^2\) With
\[ \gamma = \frac{1}{\sqrt{1 - \left( \frac{v}{c} \right)^2}} \]
(3.25)

\(^2\) Very clear if you set $c = 1$. Or,
\[ x' = \gamma (x - vt), \]
\[ t' = \gamma \left( t - \frac{vx}{c^2} \right). \]  
(3.23) (3.24)

### 3.2 Consequences of special relativity

All reference frames in this section are inertial frames.
3.2.1 Relativity of simultaneity

Two events happening in two different locations \( x_1 \) and \( x_2 \) that occur simultaneously

\[
t_1 = t_2 \quad \Rightarrow \quad \Delta t = 0
\]  
(3.26)

in the reference frame of one observer \( A \), may occur non-simultaneously in the reference frame of another observer \( B' \). Let’s move in the boosted frame \( B' \) traveling with \( v \) relative to \( A \).

\[
t'_1 = \gamma \left( t - \frac{vx_1}{c^2} \right), \quad t'_2 = \gamma \left( t - \frac{vx_2}{c^2} \right)
\]
(3.27)

\[
\Rightarrow \quad \Delta t' = t'_1 - t'_2 = \gamma \frac{v}{c^2} (x_2 - x_1) \neq 0
\]  
(3.28)

This is counter-intuitive and seems to contradict our usual notion of cause and effect. Can we change the order of two events? If \( t_2 - t_1 > 0 \), we want that

\[
0 < t'_2 - t'_1 = \gamma \left( t_2 - t_1 - \frac{v}{c^2} (x_2 - x_1) \right)
\]
(3.29)

We have to have that

\[
t_2 - t_1 > \frac{v}{c} \frac{x_2 - x_1}{c}
\]  
(3.30)

Since \( v < c \), the order of the events is kept for

\[
t_2 - t_1 \geq \frac{x_2 - x_1}{c}
\]  
(3.31)

If two events are to be causally connected, the cause can only be transmitted with finite speed \( \bar{v} \leq c \). This means

\[
t_2 - t_1 = \frac{x_2 - x_1}{\bar{v}} \geq \frac{x_2 - x_1}{c}
\]  
(3.32)
We find that cause and effect cannot be exchanged. However, the order of causally disconnected events can very well be changed.

\[ t_A - t_B = 0 \]
\[ \Delta t' > 0 \]

3.2.2 Time dilation

Let us assume that we have a clock in A which ticks every \( \Delta t \) seconds. In \( B' \) we will find the ticks at

\[ \Delta t' = \gamma \Delta t = \frac{\Delta t}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} \geq \Delta t \quad (3.33) \]

The stationary clock ticks faster! Why is the situation not symmetric? An observer in \( B' \) will have to use two clocks to compare to the for him moving clock of A.

We can now define the important quantity of **proper time** \( \Delta \tau \): it is the time measured by a clock which is at a fixed place. We can observe time dilation in (almost) everyday experiments like the decay of unstable particles, like the muons generated in the atmosphere by cosmic radiation. See Ex1!

3.2.3 Lorentz contraction

Suppose we measure a length as \( \Delta x = l \) in A. How did we perform the measurement? We put a standard ruler on the distance and compare the endpoints **simultaneously** \( \Delta t = 0 \). A measurement of the length in \( B' \) moving with \( v \) relative to A is characterized by \( \Delta t' = 0 \). The positions of the endpoints in \( B' \) are

\[ \Delta x' = \gamma (\Delta x - v \Delta t) \quad (3.34) \]
where $\Delta t = t_2 - t_1 = \gamma \Delta t' > \Delta t'$. We have to measure simultaneously in $B'$:

$$t_1 - \frac{v}{c^2} x_1 \equiv t_2 - \frac{v}{c^2} x_2$$

(3.35)

Therefore

$$t_1 - t_2 = \frac{v}{c^2} (x_1 - x_2)$$

(3.36)

Plugging back

$$l' = x'_1 - x'_2 = \gamma \left( x_1 - x_2 - \frac{v^2}{c^2} (x_1 - x_2) \right)$$

(3.37)

which means

$$l' = l \sqrt{1 - \frac{v^2}{c^2}}$$

(3.38)

A stick of length $l$ at rest in $A$ will appear contracted in $B'$ by a factor $\sqrt{1 - \frac{v^2}{c^2}} < 1$. Crucial for the length measurement is the requirement of simultaneity.

3.2.4 Addition of velocities

We can now derive the correct law of addition of velocities. Consider again the simple case of an object moving in the $x$ direction. The Lorentz transformation reads

$$x' = \frac{1}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} (x + vt),$$

(3.39)

$$t' = \frac{1}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} \left( t + \frac{vx}{c^2} \right).$$

(3.40)
Let the velocity seen in A be $u = \frac{dx}{dt}$ and in $B'$ be $u' = \frac{dx'}{dt'}$. Then dividing $dx'$ by $dt'$. We obtain

$$u' = \frac{dx'}{dt'} = v \frac{dt}{dt'} + \frac{dx}{dt'} = \frac{v + \frac{dx}{dt}}{1 + \frac{v dx}{c^2}} = \frac{v + u}{1 + \frac{vu}{c^2}}$$

(3.41)

Instead of the Galilean $u' = u + v$, the correct law contains a crucial denominator

$$u' = \frac{u + v}{1 + \frac{vu}{c^2}}$$

(3.42)

This function is remarkable: it is symmetric under $v \leftrightarrow u$. If the object is slowly moving $u \ll c$, then $u' \approx u + v$, in accordance with everyday intuition. But if a particle happens to be moving at the speed of light (e.g. a photon), $u = c$, then $u' = \frac{c + v}{1 + \frac{vu}{c^2}} = c$.

### 3.3 Minkowski space

In Euclidean space, the invariance of the combination

$$dl^2 = dx^2 + dy^2 + dz^2$$

(3.43)

allows us to define $dl$ as the distance between two points. Two observers whose coordinate systems are rotated to each other measure the same distance.\(^3\)

With profound insight, Minkowski realized\(^4\) that the invariance of the combination

$$ds^2 = dt^2 - dx^2 - dy^2 - dz^2$$

(3.44)

allows us to define $ds^2$ as the distance between two points in space-time.
allows us to talk about **distance in space-time**. This invariance
determines the Lorentz transformation\(^5\). Similar to the case of
rotations, two observers in uniform motion relative to each other can
now agree on the space-time distance between two points.

We say that the separation between two nearby points in space-
time is

\[
\begin{align*}
ds^2 &> 0 \quad dt^2 > dx^2 + dy^2 + dz^2 \quad \text{timelike} \\
ds^2 &= 0 \quad dt^2 = dx^2 + dy^2 + dz^2 \quad \text{lightlike} \\
ds^2 < 0 \quad dt^2 < dx^2 + dy^2 + dz^2 \quad \text{spacelike}
\end{align*}
\]

The first two are causally connected. Light-like is also sometimes
called 'null'.

We see that we can identify \(ds\) with a **proper-time** intervall \(d\tau\)
since it is the time measured by a resting particle \(ds^2 = dt^2\). Since
\(ds^2\) is Lorentz-invariant, we can determine it in any inertial system
by calculating \(ds^2 = dt^2 - (dx^2 + dy^2 + dz^2)\), or

\[
d\tau^2 = ds^2 = dt^2 \left(1 - \frac{dx^2 + dy^2 + dz^2}{dt^2}\right) \\
= dt^2 \left(1 - v^2(t)\right)
\]

which is exactly the already discussed time dilation.

For an arbitrarily moving particle \(x(t)\) with a velocity \(v(t) =
dx(t)/dt\), we can calculate

\[
\tau_B - \tau_A = \int_{\tau_A}^{\tau_B} dt = \int_{t_A}^{t_B} dt \sqrt{1 - v^2(t)} \leq t_B - t_A
\]

### 3.3.1 Fourvectors

We define space-time points \(X\) using the components of a four-vector

\[
x^\mu \equiv \begin{pmatrix} x^0 \\
x^1 \\
x^2 \\
x^3 \end{pmatrix} \equiv \begin{pmatrix} ct \\
x \\
y \\
z \end{pmatrix}
\]

The components are defined relative to basis vectors \(e_\mu\)

\[
X = x^\mu e_\mu = \sum_{\mu=0}^3 x^\mu e_\mu
\]

where we have used the **Einstein summation convention**, when
an index variable appears twice in a single term, it implies summation
of that term over all the values of the index. In the relativistic
context, we will always require one 'upstairs' and one 'downstairs'
index. In the following, we will therefore not show the "\(\sum_{\mu=0}^3\)"
anymore.

A **Lorentz transformation** can then be written as

\[
x'^\mu = \Lambda^\mu_\nu x^\nu
\]
where we have moved one index to the right to make obvious, which one is the first and which one is the second index of the matrix. Since we have only changed the coordinates and the physical vectors has to remain the same, we know
\[ X = x^{\mu} e_{\mu} = x'^{\mu} e'_{\mu} = \Lambda^\mu_{\nu} x^{\nu} e'_{\mu} \quad (3.51) \]
we find for the basis vectors
\[ e_{\nu} = \Lambda^\nu_{\mu} e'_{\mu} \quad (3.52) \]
and a Lorentz transformation along the \( x \)-direction would look like
\[
\begin{pmatrix}
\gamma & \gamma \beta & 0 & 0 \\
\gamma \beta & \gamma & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\quad (3.53)
\]
with \( \gamma = 1/\sqrt{1 - v^2/c^2} \) and \( \beta = v/c \). This is called a boost. We can of course also parametrize ordinary rotations within a Lorentz transformation, say a rotation around the \( z \)-Axis
\[
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & \cos \theta & -\sin \theta & 0 \\
0 & \sin \theta & \cos \theta & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\quad (3.54)
\]
3.3.2 Minkowski Metric
Lorentz transformations leave the relativistic distance \( ds^2 \) invariant. This defines a metric of the Minkowski-space
\[
ds^2 = dt^2 - dx^2 - dy^2 - dz^2 = \eta_{\mu\nu} dx^\mu dx^\nu \quad (3.55)
\]
with
\[
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 \\
0 & 0 & -1 & 0 \\
0 & 0 & 0 & -1
\end{pmatrix}
\quad (3.56)
\]
The invariance of \( ds^2 \) then implies
\[
ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu = \eta_{\mu\nu} dx'^\mu dx'^\nu = \eta_{\mu\nu} \Lambda^\mu_{\alpha} \Lambda^\nu_{\beta} dx^\alpha dx^\beta \quad (3.57)
\]
and we derive the condition
\[
\eta_{\alpha\beta} = \eta_{\mu\nu} \Lambda^\mu_{\alpha} \Lambda^\nu_{\beta} \quad (3.58)
\]
which we will use extensively later. This defines the set of all Lorentz transformations.\(^6\)

\(^6\) To get the representations of the Lorentz group \( SO(1,3) \).
3.3.3 Discrete transformations

Space-time distances as measured by the Minkowski metric, or Minkowski scalar products like

\[ V_\mu W^\mu \] (3.59)

are left invariant not only by transformations which can be continuously connected to the identity, but also by discrete transformations like parity

\[ P : (t, x, y, z) \longrightarrow (t, -x, -y, -z) \] (3.60)

or time reversal

\[ T : (t, x, y, z) \longrightarrow (-t, x, y, z) \] (3.61)

Parity and time reversal are special because they cannot be written as a product of rotations and boosts. We can write them as

\[
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 \\
0 & 0 & -1 & 0 \\
0 & 0 & 0 & -1
\end{pmatrix}, \quad
\begin{pmatrix}
-1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\] (3.62)

As you will learn in the discussion of the Standard Model of particle physics, neither symmetry is conserved in nature!

3.3.4 Motion of a free particle

Newton enunciated that in the absence of external forces, a particle will maintain a constant velocity. Hence the equation of motion

\[
\frac{d^2 X}{dt^2} = 0
\] (3.63)

or in other words \( \frac{dX}{dt} \) stays unchanged.

How does a free particle move in special relativity? Again, the answer had to be enunciated (or guessed) by Einstein. He knew that it must reduce to Newton’s equation for a slowly moving particle. Also, all inertial frame observers have to agree that the particle is freely moving. Both requirements are satisfied by

\[
\frac{d^2 X^\mu}{d\tau^2} = 0
\] (3.64)

where \( X^\mu = X^\mu(\tau) \) denotes the location of a particle with \( \tau \) being the proper time. In another inertial frame \( X'^\mu = A^\mu_\nu X^\nu \) and so

\[
\frac{d^2 X'^\mu}{d\tau'^2} = A^\mu_\nu \frac{d^2 X^\nu}{d\tau'^2} = 0
\] (3.65)

and that \( d\tau \) is the same for both observers. This motivates the introduction of the four-velocity

\[
w^\mu \equiv \frac{dx^\mu}{d\tau}
\] (3.66)
which because of \( \frac{dt}{d\tau} = \gamma(v) \) can be written in terms of the usual three-velocity \( v \) as

\[
\frac{dx^\mu}{dt} = \gamma(v) \left( \frac{c}{v} \right)
\]

(3.67)

The (Lorentz) square of the four-velocity is \(^9\)

\[
u^2 = u_\mu u^\mu = \gamma(v)^2 (c^2 - v^2) = c^2
\]

(3.68)

### 3.4 Indices upstairs and downstairs

For now \( \eta_{\mu\nu} \) is the only object with lower indices. When we want to sum over indices, the rule is that we multiply by \( \eta_{\mu\nu} \) and invoke the Einstein summation convention, e.g. for two vectors \( p^\mu q^\nu \)

\[
p \cdot q \equiv \eta_{\mu\nu} p^\mu q^\nu
\]

(3.69)

To save ourselves from constantly writing the metric \( \eta_{\mu\nu} \), we define, when we are given a vector \( p^\mu \) a vector with lower index

\[
p_\nu \equiv \eta_{\mu\nu} p^\mu
\]

(3.70)

Or in other words,

\[
p^\mu = (p^0, \mathbf{p}), \quad \text{and} \quad p_\mu = (p^0, -\mathbf{p})
\]

(3.71)

and therefore

\[
p \cdot q = p_\mu q^\mu = p^0 q^0 - \mathbf{p} \cdot \mathbf{q}
\]

(3.72)

I hope you are not confused by this trivial act of notational sloth and the fact that we have "two kinds of vectors", contravariant vectors \( p^\mu \) and covariant vectors \( p_\mu \). There is nothing profound going on here. Just a convenient notation.\(^{10}\)

Indeed, we can also define an inverse for \( \eta_{\mu\nu} \), \((\eta_{\mu\nu})^{-1} = \eta^{\mu\nu}\)

\[
\eta_{\mu\nu} \eta^{\nu\lambda} = \delta^\lambda_\mu
\]

(3.73)

As you can see it is numerically the same matrix\(^{11}\) and we will use the same symbol \( \eta \) with two covariant indices: \( \eta_{\mu\nu} \). From this we can see there is no such thing as \( \delta^{\mu\nu} \) or \( \delta_{\mu\nu} \), neither does \( \eta^\mu_\nu \) exist!

#### 3.4.1 A convenient notation and derivatives

It follows that the short-hand \( \partial_\mu \) for \( \frac{\partial}{\partial x^\mu} \) has to carry a lower index because

\[
\partial_\mu x^\nu = \frac{\partial x^\nu}{\partial x^\mu} = \delta^\nu_\mu
\]

(3.74)

We will use this fact repeatedly. You can also argue that a variation of a Lorentz invariant scalar field \( \phi(x^\mu) \) should still be a scalar

\[
\delta \phi = \frac{\partial \phi}{\partial x^\mu} \delta x^\mu
\]

(3.75)
and therefore
\[
\partial_{\mu} = \frac{\partial}{\partial x^\mu} = \left( \frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)
\] (3.76)
and
\[
\partial^\mu = \frac{\partial}{\partial x_\mu} = \left( \frac{\partial}{\partial t}, -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z} \right)
\] (3.77)

A useful mnemonic is to think of \(x^\mu\) carrying an upper index but since in \(\frac{\partial}{\partial x_\mu}\) it appears in the denominator, it should carry a lower index.

We will also use
\[
\partial_\mu A^\mu = \partial_0 A^0 + \partial_j A^j
\] (3.78)
and
\[
\partial_\mu \partial^\mu = \frac{\partial^2}{\partial t^2} - \nabla^2 = \Box
\] (3.79)

Finally, we will make use of the completely anti-symmetric tensor \(\epsilon^{\mu\nu\alpha\beta}\), known as the *Levi-Civita tensor*. It is defined as
\[
\epsilon^{\mu\nu\alpha\beta} = \begin{cases} 
1, & \text{if } \mu\nu\alpha\beta \text{ is an even permutation of } (0,1,2,3) \\
-1, & \text{if } \mu\nu\alpha\beta \text{ is an odd permutation of } (0,1,2,3) \\
0, & \text{if } \mu\nu\alpha\beta \text{ is not a permutation of } (0,1,2,3)
\end{cases}
\] (3.80)
So e.g. \(\epsilon^{0123} = -\epsilon^{1023} = 1\) and \(\epsilon^{0223} = 0\).

3.4.2 The appearance of anti-matter

An actual physics process (event \(A\)), a proton turns into a neutron by emitting a pion, the \(\pi^+\), which by conservation of charge necessarily carries positive charge. In event \(B\), the \(\pi^+\) is absorbed which turns a neighboring neutron into a proton. This generates an attraction between the proton and the neutron and is an effective description of the strong force.

The same process as seen by two observers. Since the pion propagates over a spacelike interval, it is possible to see a temporal order in which \(A\) occurs after \(B\). This something, by charge conservation, has to carry negative charge which implies the existence of anti-matter.

3.5 Relativistic mechanics

We define the four-momentum of a particle of mass \(m\) with
\[
p^\mu \equiv m u^\mu
\] (3.84)
which we can write in an inertial system where the particle has the three-velocity \(v(t)\)
\[
p^\mu = \begin{pmatrix} p^0 \\ p \end{pmatrix} = m \gamma(v) \begin{pmatrix} c \\ v \end{pmatrix}
\] (3.85)
Figure 3.6: The need for antimatter. The same process as seen by two observers. Since the pion propagates over a spacelike interval, it is possible to see a temporal order in which $A$ occurs after $B$. This something $\pi^+$, by charge conservation, has to carry negative charge.

The dual space: The Minkowski metric defines a (not positive definite) norm for four vectors

$$\Delta s^2 = (\Delta X, \Delta X') = \Delta x^\mu \Delta x'^\nu (e_\mu, e_\nu)_{\eta_{\mu\nu}}$$

(3.81)

and a Lorentz-invariant scalar product for arbitrary vectors $A, B$

$$\langle A, B \rangle = a^\mu b'_\nu \eta_{\mu\nu}$$

(3.82)

Every vector space $V$ has a dual space, often denoted $V^*$, which is simply the set of all linear maps from the vector space into the real (or complex) numbers. We can find a basis $e^\mu$ for the dual space with

$$e^\mu[e_\nu] = \delta^\mu_\nu$$

(3.83)

where $e^\mu[..]$ takes a vector to produce a number.

The metric gives us a canonical isomorphism between the space of covariant vectors $V$ and its dual space $V^*$, allowing us to be sloppy and conflate the two. More general objects with $n$ upper indices and $m$ lower indices can be thought of as living in a tensor product of copies of $V$ and $V^*$, i.e. $V^{\otimes m} \otimes V^*^{\otimes n}$.

For us contra-variant and co-variant vectors are just a convenient notation but if you are a mathematician-want-to-be you can read about this in more profound books, or here [https://en.wikipedia.org/wiki/Dual_space](https://en.wikipedia.org/wiki/Dual_space).

The three-momentum is therefore

$$p = m\gamma(v) v$$

(3.86)

This coincides with the usual definition of momentum if we introduce a velocity dependent mass $m(v) = \gamma(v)m$. We call $m = m(0)$ the rest mass.

The relativistic energy$^{12}$ $E = cp^0$ is proportional to the 0-component

$$E = cp^0 = m\gamma(v)c^2 = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} \approx mc^2 \left(1 + \frac{1}{2} \frac{v^2}{c^2} + \ldots \right)$$

(3.87)

where a new term, the rest energy, appeared. We find for the square of the momentum

$$p^2 = p_\mu p^\mu = (p^0)^2 - p^2 = m^2 c^2$$

(3.88)

where we have used $u_\mu u^\mu = c^2$.
3.5.1 $E = mc^2$

The energy and the momentum of a particle of mass $m$ form a four-vector

$$p^\mu = \begin{pmatrix} E \\ p \end{pmatrix}$$

whose square is

$$p^2 = E^2 - p^2 = m^2$$

(3.90)

This is the relativistic Pythagoras which defines a Lorentz-invariant relationship between the energy, the rest-mass and the momentum of a particle:

$$E^2 = m^2 + p^2$$

(3.91)

3.5.2 Application: weak decay of a pion

A charged $\pi^\pm$ can decay to a muon $\mu^\pm$ and a muon anti-neutrino $\nu_\mu$. The muon has about $3/4$ the mass of a pion and the neutrino is almost massless. We can calculate the energy and the velocity of the decay products using the four-momentum conservation

$$p^\mu_{\text{ini}} = p^\mu_{\text{final}}$$

(3.92)

which is

$$\begin{pmatrix} E_{\pi^+}^\mu \\ p_{\pi^+}^\mu \end{pmatrix} = \begin{pmatrix} E_{\mu^+}^\mu + m_{\mu^+}^2 \\ p_{\mu^+}^\mu + p_{\nu_\mu} \end{pmatrix} = \begin{pmatrix} E_{\text{final}}^\mu \\ p_{\text{final}}^\mu \end{pmatrix}$$

(3.93)

In the rest-system of the pion we derive from energy conservation

$$E_{\pi^+} = m_{\pi^+} = E_{\mu^+} + E_{\nu_\mu}$$

(3.94)

The almost mass-less neutrino has $E_{\nu_\mu} = |p_{\nu_\mu}|$ and because the initial three-momentum was zero in the rest system of the pion, we get

$$|p_{\nu_\mu}| = |p_{\mu^+}| = \frac{m_{\mu^+}^2 - m_{\nu}^2}{2m_{\pi^+}}$$

(3.95)

For the decay to an electron and an electron anti-neutrino, the electron mass is negligible: we find the decay products back-to-back and the rest-mass has been converted entirely into kinetic energy with $|p_{\nu}| = |p_{\mu^+}| = m_{\pi^+}/2$. 

Figure 3.7: Feynman diagram of a $\pi^+$ decay. $m_{\pi^\pm} = 139.57$ MeV and $m_{\mu^\pm} = 105.658$ MeV.
4

Classical Field Theory

We will first introduce various aspects of classical fields. We will return to the subject at several later stages when we need new concepts.

4.1 Dynamics of fields

In QM, observables are constructed out of $q$’s and $p$’s. Classical fields are defined at each point in space-time.

$$\phi_a(x^\mu)$$ (4.1)

Thus we are dealing with a system with infinitely many degrees of freedom, at least one for each point $x$ in space. Notice, that the concept of position has been demoted from a dynamical variable in particle mechanics to a mere label in field theory.

4.1.1 Example: the electro-magnetic field

You are very familiar with $E(x, t)$ and $B(x, t)$ from classical electrodynamics. In a more advanced treatment of electromagnetism, we derive the spatial 3-vectors from a single 4-component field

$$A^\mu(x^\mu) = \begin{pmatrix} \phi \\ A \end{pmatrix}$$ (4.2)

where the vector potential $A^\mu$ is a vector in space-time. The original electric and magnetic fields are related by

$$E = -\nabla \phi - \frac{\partial A}{\partial t}, \quad \text{and} \quad B = \nabla \times A$$ (4.3)

which have two of Maxwell’s equations as an immediate consequence

$$\nabla \cdot B = 0, \quad \text{and} \quad \frac{\partial B}{\partial t} = -\nabla \times E$$ (4.4)

4.1.2 The Lagrangian density

The dynamics of fields is determined by a Lagrangian which is a function of $\phi(x, t), \dot{\phi}(x, t), \nabla \phi(x, t)$. All the systems in this course will have the form

$$L(t) = \int d^4x \, L(\phi_a, \partial_\mu \phi_a)$$ (4.5)
where the $\mathcal{L}$ is the Lagrangian density$^2$. The action is

$$S = \int_{t_1}^{t_2} dt \int d^3x \mathcal{L} \quad (4.6)$$

Remember that in classical mechanics, $L$ depends on $q$ and $\dot{q}$. In field theory we analogously restrict to Lagrangians depending on $\phi$ and $\dot{\phi}$ and not on $\ddot{\phi}$. We could in principle write Lagrangian's which depend on $\nabla \phi$, $\nabla^2 \phi$, $\nabla^3 \phi$ and so on. But anticipating Lorentz-invariance, we will only consider $\nabla \phi$ and no higher derivatives. Also we will not include an explicit $x^\mu$ dependence, all such dependence comes only implicitly through $\phi$ and its derivatives.

In the following, we will use functional derivatives. A functional $F[\phi]$ for us is a map from a vector space of functions into the real numbers, like e.g. the action $S[\phi]$. A functional derivative is denoted by $\frac{\delta F[\phi]}{\delta \phi(x')}$ and it is defined as the linear term in an expansion

$$F[\phi + \eta] = F[\phi] + \int dx' \frac{\delta F[\phi]}{\delta \phi(x')} \eta(x') + \ldots \quad (4.7)$$

or

$$\int dx \frac{\delta F[\phi]}{\delta \phi}(x) \eta(x) = \lim_{\varepsilon \to 0} \frac{F[\phi + \varepsilon \eta] - F[\phi]}{\varepsilon} = \left[ \frac{d}{d\varepsilon} F[\phi + \varepsilon \eta] \right]_{\varepsilon = 0}, \quad (4.8)$$

where $\eta$ is an arbitrary function. In particular, for the functional $F[\phi] = \phi(x)$, we find

$$\frac{\delta F[\phi]}{\delta \phi}(x) = \frac{\delta \phi(x)}{\delta \phi(x')} = \delta(x - x') \quad (4.10)$$

The quantity $\varepsilon \eta$ is called the variation of $\phi$. The differential (or variation or first variation) of the functional $F[\phi]$ is

$$\delta_i F(\phi) = \int dx \frac{\delta F}{\delta \phi_i}(x) \eta(x) \quad (4.11)$$

This is similar in form to the total differential of a function. You should familiarize yourself with the mechanics of functional derivatives on the next exercise sheet.

We can derive the equations of motion by the principle of least action. We vary the field at each space-time point, keeping the end-points fixed and requiring

$$\delta S = 0 \quad (4.12)$$

with

$$\delta S = \int d^4x \left[ \frac{\partial \mathcal{L}}{\partial \delta \phi_a} \delta \phi_a + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} \delta (\partial_\mu \phi_a) \right]$$

$$= \int d^4x \left\{ \left[ \frac{\partial \mathcal{L}}{\partial \phi_a} - \partial_\mu \left( \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} \right) \right] \delta \phi_a + \partial_\mu \left( \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} \delta \phi_a \right) \right\} \quad (4.13)$$

The last term is a total derivative and therefore its integral only depends on the field values at the spatial and temporal boundaries. It vanishes for any $\delta \phi_a$ that decays at spatial infinity and obeys

$$\delta \phi_a(x, t_1) = \delta \phi_a(x, t_2) = 0 \quad (4.14)$$

Note, we will always make the assumption that we can drop such total derivatives from the Lagrangian. It lets us integrate by parts without consequence. We will therefore equate

$$\phi \partial_\mu \psi = - (\partial_\mu \phi) \psi$$

in a Lagrangian.
Requiring \( \delta S = 0 \) then yields the **Euler-Lagrange** equations of motions for the fields

\[
\frac{\partial \mathcal{L}}{\partial \phi_a} - \partial_\mu \left( \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} \right) = 0 \tag{4.15}
\]

### 4.1.3 Example 1: the Klein-Gordon equation

Consider a real scalar field \( \phi(x^\mu) \) and the Lagrangian

\[
\mathcal{L} = \frac{1}{2} \eta_{\mu\nu} \partial_\mu \phi \partial^\nu \phi - \frac{1}{2} m^2 \phi^2 \tag{4.16}
\]

Comparison with the usual structure \( \mathcal{L} = T - V \) we see, that we have a **kinetic energy**

\[
T = \int d^3x \frac{1}{2} (\dot{\phi})^2 \tag{4.18}
\]

and the **potential energy** of the field

\[
V = \int d^3x \left[ \frac{1}{2} (\nabla \phi)^2 + \frac{1}{2} m^2 \phi^2 \right] \tag{4.19}
\]

where the first term is called gradient energy, while we usually call the second term the "potential". Using

\[
\frac{\partial \mathcal{L}}{\partial \phi} = -m^2 \phi, \quad \text{and} \quad \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} = \partial^\mu \phi \tag{4.20}
\]

we get the Euler-Lagrange equation for the **Klein-Gordon** field

\[
\partial_\mu \partial^\mu \phi + m^2 \phi = 0 \tag{4.21}
\]

If we add an arbitrary potential \( V(\phi) \) we can generalize the Klein-Gordon equation to

\[
\Box \phi + \frac{\partial V}{\partial \phi} = 0 \tag{4.22}
\]

where we have used \( \Box = \partial_\mu \partial^\mu \) for the Laplacian in Minkowski space.

### 4.1.4 Example 2: first order Lagrangians

Alternatively we can write a Lagrangian which is linear in time derivatives. Take a complex scalar field \( \psi \) whose dynamics is defined by a real Lagrangian

\[
\mathcal{L} = \frac{i}{2} (\psi^* \partial_t \psi - (\partial_t \psi^*) \psi) - \frac{1}{2m} \nabla \psi^* \cdot \nabla \psi - \frac{m \omega^2}{2} \psi \tag{4.23}
\]

We derive the EOMs by treating the \( \psi \) and \( \psi^* \) as independent variables

\[
\frac{\partial \mathcal{L}}{\partial \psi^*} = \frac{i}{2} \partial_t \psi - \frac{m \omega^2}{2} \psi \tag{4.24}
\]

\[
\frac{\partial \mathcal{L}}{\partial (\partial_t \psi^*)} = -\frac{i}{2} \psi \tag{4.25}
\]

\[
\frac{\partial \mathcal{L}}{\partial (\nabla \psi^*)} = -\frac{1}{2m} \nabla \psi \tag{4.26}
\]
from which we obtain

\[ i \frac{\partial \psi}{\partial t} = -\frac{1}{2m} \nabla^2 \psi + \frac{m \omega^2}{2} \psi \]  

This looks like the Schrödinger equation except for the harmonic potential, but its interpretation is very different. The field \( \psi \) is a classical field and we have no probability interpretation attached to it.

The plane which gives the initial conditions that determine the future (and the past) uniquely is called a \textbf{Cauchy surface}. It differs between the two examples above. If \( \mathcal{L} \propto (\partial_t \phi)^2 \), we must specify both \( \phi \) and \( \partial_t \phi \) to fix the future evolution. For \( \mathcal{L} \propto \psi^* \partial_t \psi \), we only need \( \psi \) and \( \psi^* \).

\subsection{Example 3: the Maxwell equations}

The Maxwell equations without sources are given by the Lagrangian

\[ \mathcal{L} = -\frac{1}{2} (\partial_\mu A_\nu)(\partial^\mu A^\nu) + \frac{1}{2} (\partial_\mu A^\mu)^2 \]  

Note that the minus signs are crucial to ensure positive kinetic terms for the \( A_i \),

\[ \mathcal{L} = \frac{1}{2} (\partial_\nu A^\nu) + \frac{1}{2} (\nabla A^0)^2 - \frac{1}{2} (1 - \delta_{ij}) (\partial_i A^j)^2 \]  

Observe, that there is no kinetic term for \( A_0 \) (no \( \mathcal{L} \propto (\partial_t A^0)^2 \)), which will have important consequences in QED. With

\[ \frac{\partial \mathcal{L}}{\partial (\partial_\mu A_\nu)} = -\partial^\mu A^\nu + (\partial_\alpha A^\alpha) \eta^{\mu\nu} \]  

we get the Euler-Lagrange-equations

\[ \partial^\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu A_\nu)} = -\partial^2 A^\nu + \partial^\nu (\partial_\alpha A^\alpha) = -\partial_\mu (\partial^\mu A^\nu - \partial^\nu A^\mu) \]  

\[ \equiv -\partial_\mu F^{\mu\nu} \]  

where the \textbf{field strength} is defined as\(^3\)

\[ F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \]

The Maxwell Lagrangian therefore can be written in the compact form

\[ \mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \]

where we’ve used an integration by parts.

\subsection{Locality}

All examples above are \textbf{local} Lagrangians. We have not written any terms that look like

\[ L = \int d^3x d^3y \phi(x) \phi(y) \]
which would couple a field at \( x \) to a field at \( y \). The closest we get is a coupling between \( \phi(x) \) and \( \phi(x + \delta x) \) through the gradient term \((\nabla \phi(x))^2\). This property of locality is, as far as we know, a key feature of all theories of Nature. Locality has a deep connection to unitarity and one finds that non-local theories quickly violate unitarity.

### 4.1.7 Lorentz invariance

Lorentz transformations have a representation on the fields. The simplest example is the scalar field which, under the Lorentz transformation \( x^\mu \rightarrow \Lambda^\mu_\nu x^\nu \), transforms as

\[
\phi(x) \rightarrow \phi'(x) = \phi(\Lambda^{-1} x)
\]

There is an inverse \( \Lambda^{-1} \) in the argument because we are dealing with an active transformation in which we truly shift the field.

We call a theory Lorentz invariant if \( \phi(x) \) solves the EOMs then also \( \phi(\Lambda^{-1} x) \) solves the EOMs. We can ensure that this holds by requiring the action to be Lorentz invariant. Let’s look at our three examples.

1. **Klein-Gordon equation**: For our real scalar field we have \( \phi(x) \rightarrow \phi'(x) = \phi(\Lambda^{-1} x) = \phi(y) \). The contraction of derivatives is obviously invariant using Eq. (3.58). The potential terms transform similarly, with \( \phi^2(x) \rightarrow \phi^2(y) \). Putting this all together, we find that the action is indeed invariant under Lorentz transformations,

\[
S = \int d^4 x \mathcal{L}(x) = \int d^4 x \left[ \partial_{\mu} \phi(x) \partial^{\mu} \phi(x) - m^2 \phi^2(x) \right] \rightarrow \int d^4 y \left[ \partial'_{\mu} \phi(y) \partial'^{\mu} \phi(y) - m^2 \phi^2(y) \right] = \int d^4 y \mathcal{L}(y) = S
\]

where we’ve denoted \( \partial_{\mu} = \frac{\partial}{\partial x^\mu} \) and used the fact that the Jacobian is trivial

\[
d^4 x \rightarrow \det(\Lambda) d^4 y = d^4 y
\]

with \(^4\)

\[
\det(\Lambda) = 1
\]

for Lorentz transformations connected to the identity, which we are using for now.

2. **First order equation**: The theory is not Lorentz invariant since space and time are not on the same footing. In practice, it’s easy to see if the action is Lorentz invariant: just make sure all the Lorentz indices \( \mu, \nu, \ldots \) are contracted with Lorentz invariant objects, such as the metric \( \eta_{\mu\nu} \).

3. **Maxwell’s equations**: Under a Lorentz transformation we obtain

\[
A^\mu(x) \rightarrow \Lambda^\mu_\nu A^\nu(\Lambda^{-1} x)
\]

Again, effective field theories are special and can describe the correct physics until some scale while being non-local.

An example for an active transformation: start with a temperature field \( \phi(x) \) which has a hotspot at \( x = (1,0,0) \). After a rotation

\[
x \rightarrow Rx
\]

The new field \( \phi'(x) \) will have the hotspot at \( x = (0,1,0) \). If we want to express \( \phi'(x) \) in terms of the old field \( \phi \), we need to place ourselves at \( x = (0,1,0) \) and ask ourselves what the old field looked like where we’ve come from at

\[
R^{-1} x = R^{-1}(0,1,0) = (1,0,0)
\]

See next comment on passive transformations.

Continuing with the temperature example: This \( R^{-1} \) is the reason for the inverse transformation in the formula.

If we were instead dealing with a passive transformation in which we relabel our choice of coordinates, we would have instead

\[
\phi(x) \rightarrow \phi'(x) = \phi(\Lambda x)
\]

\[^4\] Ex: Show that

\[
|\det(\Lambda)| = 1
\]

using Eq. (3.58).
and you can see that Maxwell’s Lagrangian is indeed invariant.
Historically of course, it was Maxwell’s Lagrangian which first
show the concept of Lorentz invariance.

4.2 Symmetries: Noether’s Theorem and conserved currents

Symmetries in field theory are probably even more important than
in particle mechanics. Every additional symmetry constraints and
reduces our freedom to pick terms and their coefficients in the
Lagrangian. There are Lorentz symmetries, internal symmetries (like
$U(1)$ or $SU(3)$ in QCD), gauge redundancies, supersymmetry and so
on.

They are the real point of the Lagrangian formalism: the La-
grangian provides a natural framework for the quantum mechanical
implementation of symmetry principles. Noether’s theorem is useful
because it allows you to make exact statements about the solutions
of a theory without solving it explicitly. Since in quantum field the-
ory we won’t be able to solve almost anything exactly, symmetry
arguments will be extremely important!

**Noether’s theorem:** Every continuous symmetry of the action
gives rise to a conserved current $j_\mu(x)$ such that the equations of
motion imply

$$\partial_\mu j^\mu(x) = 0$$  \hspace{0.5cm} (4.36)
which in components is $\partial_\mu j^0 + \nabla \cdot j = 0$. We will now prove this.

Since we are assuming a continuous symmetry, we may prove it by
working infinitesimally. The transformation

$$\delta \phi_a(x) = X_a(x)$$ \hspace{0.5cm} (4.37)
is a symmetry of the action. This means that the Lagrangian
changes at most by a total derivative

$$\delta \mathcal{L} \bigg|_{\text{EOM}} = \partial_\mu B^\mu$$  \hspace{0.5cm} (4.38)
For an arbitrary variation $\delta \phi_a(x)$, we know that

$$\delta \mathcal{L} = \left[ \frac{\partial \mathcal{L}}{\partial \phi_a} - \partial_\mu \left( \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} \right) \right] \delta \phi_a \delta_\mu + \partial_\mu \left( \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} \right) \delta \phi_a$$  \hspace{0.5cm} (4.39)
When the equations of motion are satisfied, the first term [...] vanishes! We then have

$$\delta \mathcal{L} \bigg|_{\text{EOM}} = \partial_\mu \left( \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} \right) \delta \phi_a$$ \hspace{0.5cm} (4.40)
For a transformation which is a symmetry of the action, we know
however that $\delta X_a(x) \mathcal{L} = \partial_\mu B^\mu$. If we restrict ourselves to variations
that satisfy the EOMs, we obtain for the transformations which leave
the action invariant

$$\delta X_a(x) \mathcal{L} \bigg|_{\text{EOM}} = \partial_\mu \left( \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a)} X_a(x) \right) = \partial_\mu B^\mu$$ \hspace{0.5cm} (4.41)
or
\[ \partial_\mu \left( \frac{\partial L}{\partial (\partial_\mu \phi_a)} X_a(x) - B^\mu \right) = 0 \] (4.42)

This proves the theorem and we find
\[ \partial_\mu j^\mu(x) = 0 \quad \text{with} \quad j^\mu(x) = \frac{\partial L}{\partial (\partial_\mu \phi_a)} X_a(x) - B^\mu \] (4.43)

### 4.2.1 Conserved currents

A conserved current implies a conserved charge \( Q \) with
\[ Q \equiv \int d^3 x j^0 \] (4.44)

where the integration is over spatial coordinates only. We can see this by taking the time-derivative
\[ \frac{dQ}{dt} = \int d^3 x \frac{\partial j^0}{\partial t} = - \int d^3 x \nabla \cdot \mathbf{j} = 0 \] (4.45)

where we have assumed that \( \mathbf{j} \to 0 \) fast enough as \( x \to \infty \).

There is one such conserved current \( j^\mu \) and one constant of the motion \( Q \) for each independent infinitesimal symmetry transformation.

A conserved current is a stronger statement than a global conserved charge (i.e. integrated over all of \( \mathbb{R}^3 \)), since a conserved current implies a local conservation of charge! With the analysis above for a finite volume \( V \) in \( \mathbb{R}^3 \)
\[ Q_V = \int_V d^3 x j^0 \] (4.46)

we find\(^5\)
\[ \frac{dQ_V}{dt} = - \int_V d^3 x \nabla \cdot \mathbf{j} = - \int_{\partial V} dS \cdot \mathbf{j} \] (4.47)

with Stokes theorem. Any charge leaving \( V \) per time interval must be equal to the flow of three-current out of the volume. This holds in any local field theory.

### 4.2.2 Example 1: internal symmetries \( U(1) \)

A global internal symmetry only involves a transformation of the fields and acts the same at every point in space-time. A simple example can be obtained using a complex scalar field \( \psi(x) \)
\[ \psi(x) = \frac{1}{\sqrt{2}} (\phi_1(x) + i \phi_2(x)) \]

with \( \phi_i \) real fields. We assume the real Lagrangian
\[ \mathcal{L} = \partial_\mu \psi^* \partial^\mu \psi - V(\psi^* \psi) \] (4.48)
where the potential is a general polynomial in $|\psi|^2 = \psi^\ast \psi$. Instead of expanding in $\phi_1, \phi_2$ and deriving the equations of motion, we can equivalently treat $\psi^\ast, \psi$ as independent variables. We obtain the EOM

$$\partial_\mu \partial^\mu \psi + \frac{\partial V(\psi^\ast \psi)}{\partial \psi^\ast} = 0$$

(4.49)

after varying $\psi^\ast$. The Lagrangian is invariant under phase rotations of $\psi$

$$\psi(x) \to e^{i\alpha} \psi(x), \quad \text{or} \quad \delta \psi = i\alpha \psi$$

(4.50)

This is called a **global $U(1)$ symmetry**. We say the internal symmetry is $G = U(1)$. This implies a conservation law, since the Lagrangian remains invariant under this transformation $\delta L = 0$, and so $B^\mu = 0$. We obtain for the conserved current

$$j^\mu = \frac{\partial L}{\partial (\partial_\mu \phi_a)} X_a(x) = \frac{\partial L}{\partial (\partial_\mu \psi)} \delta \psi + \frac{\partial L}{\partial (\partial_\mu \psi^\ast)} \delta \psi^\ast$$

(4.51)

and find

$$j^\mu = i(\partial^\mu \psi^\ast) \psi - i\psi^\ast (\partial^\mu \psi)$$

(4.52)

Once we quantize this field theory, we will see that it can be interpreted as electric charge and implies a conservation of particle number.

### 4.2.3 Example 2: an internal $SO(N)$ and a bit on Lie groups

We can rewrite this Lagrangian using a basis of real fields

$$L = \partial_\mu \phi_a \partial^\mu \phi_a - V(\phi_a \phi_a)$$

(4.53)

We see that all the terms are polynomials of the form $\phi_a \phi_a = \phi_1^2 + \phi_2^2$, similar to $r^2 = x^2 + y^2$ which is invariant under rotations. We reproduce the complex transformation with

$$\begin{pmatrix} \phi_1(x) \\ \phi_2(x) \end{pmatrix} \to \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} \phi_1(x) \\ \phi_2(x) \end{pmatrix}$$

(4.54)

or

$$\begin{pmatrix} \delta \phi_1(x) \\ \delta \phi_2(x) \end{pmatrix} \to \begin{pmatrix} 0 & \alpha \\ -\alpha & 0 \end{pmatrix} \begin{pmatrix} \phi_1(x) \\ \phi_2(x) \end{pmatrix}$$

(4.55)

which we can write as

$$\delta \phi_i(x) = \alpha \epsilon^{ij} \phi_j(x)$$

(4.56)

with $\epsilon^{ij} = -\epsilon^{ji}$ and $\epsilon^{12} = 1$. This is a two dimensional, real rotation or $G = O(2)$. If we restrict ourselves to transformations which can be continuously connected to the identity, we would have $G = SO(2)$ which has the additional restriction that the group elements have unit determinant or

$$\det O = 1$$

(4.57)

6 Not even a total derivative.

7 We drop the constant $\alpha$ and we perform the sum over internal indices 1, 2 by including $\psi$ and $\psi^\ast$.

8 A summation over repeated indices is implied here, too.
This is called the **special orthogonal group** in 2 dimensions.\(^9\)

Infinitesimal group elements can be written as

\[
O = 1 + \alpha \varepsilon^{ij} + \ldots = 1 + i\alpha T^{ij} + \ldots
\]  

(4.58)

with \(O \in G = SO(2)\). We call the matrix \(T^{ij}\) parameterizing an infinitesimal rotation, a (or "the" since there is only one here) **generator** of the group. We could have written the transformation of \(\phi_a\) as

\[
\delta \phi_a(x) = i\alpha T^{ij} \phi_j(x)
\]

(4.59)

and find the conserved current as

\[
j^\mu = (\partial^\mu \phi_1)\phi_2 - (\partial^\mu \phi_2)\phi_1
\]

(4.60)

or in terms of the generator

\[
j^\mu = iT^{ij}(\partial^\mu \phi_i)\phi_j
\]

(4.61)

Generally, we obtain one conserved current for each infinitesimal transformation, or generator, that leaves the action invariant.

We can generalize this symmetry to \(a = 1, 2, \ldots, N\). This is a **non-Abelian symmetry group** \(G = SO(N)\). Similarly, we could sum a complex field \(\psi_a\) over an index \(a\) and build a Lagrangian with polynomials of \(\psi^*_a \psi_a\) (and derivatives thereof) to obtain a \(G = SU(N)\) invariant theory. \(SU(N)\) is complex non-abelian group and contains all special unitary transformations in \(N\) dimensions. These groups are called non-abelian, because their generators in general do not commute, e.g. for \(SO(3)\)

\[
[T^a, T^b] = i\varepsilon^{abc} T^c
\]

(4.62)

where we have suppressed the obvious matrix indices of the generators. We find 3 conserved currents:

\[
j^\mu_a = iT^{ij}_a(\partial^\mu \phi_i)\phi_j
\]

(4.63)

We will return to currents and conserved charges once we have introduced quantized fields.\(^10\)

\[4.2.4 \quad \text{Example 3: space-time translations}\]

Let us now study the case in which we transform space and time. Recall that in classical mechanics, invariance under spatial translations gives rise to the conservation of momentum, while invariance under time translations leads to the conservation of energy. Consider an infinitesimal transformation\(^11\)

\[
x^\nu \to x^\nu - \epsilon^\nu
\]

(4.64)

from which we obtain

\[
\phi_a(x) \to \phi_a(x) + \epsilon^\nu \partial_\nu \phi_a(x)
\]

(4.65)

\[\text{\footnotesize \textcopyright 2007 McGraw-Hill Education. All rights reserved.}\]

\[\text{\footnotesize 10 We will always factor } G \text{ into a direct product of semi-simple groups and discrete factors. For the conservation laws, the discrete factors are not important.}\]

\[\text{\footnotesize 11 Also called the } \textbf{Poincaré} \text{ transformations.}\]
Note the "+" sign because we are performing an active transformation. The Lagrangian itself will transform as

$$L(x) \rightarrow L(x) + \epsilon^\nu \partial_\nu L(x)$$

(4.66)

Noether’s theorem tells us that there are four conserved currents \((j^\nu)_\nu\) for each one of the four space-time translations \(\epsilon^\nu\). Together with the total derivative \((B^\nu)_\nu = \delta^\mu_\nu \epsilon^\mu L\), we find

$$\epsilon^\nu \partial_\nu L \approx T^\nu_\nu$$

(4.67)

This defines the energy-momentum-tensor \(T^\mu_\nu\). The four conserved quantities are the total energy \(E\)

$$E = \int d^3x T^{00}$$

(4.68)

and

$$P^i = \int d^3x T^{0i}$$

(4.69)

where \(P^i\) is the total physical momentum\(^{12}\) carried by the field.

For the Klein-Gordon Lagrangian, we can e.g. compute\(^{13}\)

$$T^{\mu\nu} = \partial^\mu \phi \partial^\nu \phi - \eta^{\mu\nu} L$$

(4.70)

We find for \(E\) and \(P^i\)

$$E = \int d^3x \left[ \frac{1}{2} (\partial_t \phi)^2 + \frac{1}{2} (\nabla \phi)^2 + \frac{1}{2} m^2 \phi^2 \right]$$

(4.71)

$$P^i = \int d^3x (\partial_i \phi) \partial^\phi$$

(4.72)

or with \(\partial^i = -\partial_i = -\nabla_i\), which we will use later:

$$P = -\int d^3x (\partial_t \phi) \nabla \phi$$

(4.73)

Note that the energy momentum tensor accidentally came out symmetric here, such that \(T^{\nu\mu} = T^{\mu\nu}\). This might not always be the case, but we can always find a symmetric form\(^{14}\) The trick is that we add an extra term which is anti-symmetric in the first two indices \(\Gamma^{\rho\mu\nu} = -\Gamma^{\mu\rho\nu}\)

$$\Theta^{\mu\nu} = T^{\mu\nu} + \partial_\rho \Gamma^{\rho\mu\nu}$$

because the anti-symmetry of \(\Gamma^{\rho\mu\nu}\) guarantees that \(\partial_\mu \partial_\rho \Gamma^{\rho\mu\nu} = 0\),\(^{15}\) which means that \(\partial_\mu \Theta^{\mu\nu} = 0\) still holds.

In general relativity, the metric \(g_{\mu\nu}\) is a field and we can expand it as \(g_{\mu\nu} = \eta_{\mu\nu} + \sqrt{G_N} h_{\mu\nu} + \ldots\). Once we enter this into the Einstein action and expand, we will find for the linear terms in \(h_{\mu\nu}\) a coupling \(h_{\mu\nu} \Theta^{\mu\nu}\).
4.2.5 Summary of Noether’s theorem

We summarize Noether’s theorem:

1. There is one conserved current $j^\mu$ and one constant of motion $Q$
   for each independent infinitesimal symmetry transformation.
2. The symmetry must be continuous, otherwise the $\delta \phi_a$ has no
   meaning.
3. The current is conserved on-shell, that is when the equations of
   motions are satisfied.
4. It applies to global symmetries, parametrized by numbers $\alpha$, e.g.
   $\phi \to e^{i\alpha} \phi$.\(^{16}\)

Currents will appear as sources for vector fields, e.g. appearing in
the Lagrangian as

$$\mathcal{L} = \ldots - j_\mu(x) A^\mu(x). \quad (4.74)$$

This can be an explicit external current such as the charge current or
just a formal place-holder for the composite expression

$$j^\mu = i(\partial^\mu \psi^*) \psi - i\psi^*(\partial^\mu \psi) \quad (4.75)$$

We will discuss this issue in more detail once we tackle the issue of
gauge invariance.

4.3 Coulomb’s law and the propagator method

Let us do some calculations to better understand aspects of classical
field theory and introduce Fourier transformation techniques along
the way. We will derive Coulomb’s law from relativistic electro-
dynamics.

A charge $e$ at the origin can be represented by the current $j^\mu$

$$j^\mu = \begin{pmatrix} \rho(x) \\ \psi^i(x) \end{pmatrix} = \begin{pmatrix} e \delta^3(x) \\ 0 \end{pmatrix} \quad (4.76)$$

Adding the source term to the Lagrangian, we get

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - j_\mu(x) A^\mu(x). \quad (4.77)$$

The Euler-Lagrange equations are

$$0 = \frac{\partial \mathcal{L}}{\partial A^\nu} - \partial_\mu \left( \frac{\partial \mathcal{L}}{\partial (\partial_\nu A^\mu)} \right) = -j_\mu + \partial^2 A^\nu - \partial^\nu (\partial_\alpha A^\alpha) \quad (4.78)$$

$$= -j_\mu + \partial_\mu F^{\mu\nu} \quad (4.79)$$

where we have used Eq. (4.31). As you may remember from your
electro-dynamics lectures, we can freely choose a gauge for $A_\mu$ and
we will use Lorenz gauge $\partial_\alpha A^\alpha = 0$ to obtain

$$\Box A^\nu = j_\mu \quad (4.80)$$
which has the formal solution

\[ A^\nu = \frac{1}{\Box} j^\mu \]  

(4.81)

where \(1/\Box\) is just the inverse of \(\Box\) which we will define more precisely later. This is a very common procedure in quantum field theory: we say that the field \(A^\nu\) is determined by the source \(j^\nu\) after it propagates with the propagator

\[ \Pi_A = \frac{1}{\Box} \]  

(4.82)

We will discuss propagators in great detail. For our example above we have to solve the following equations

\[ A_i(x) = 0 \]  

(4.83)

\[ A_0(x) = \frac{e}{\Box} \delta^3(x) \]  

(4.84)

The homogeneous solutions \(\Box A_\nu = 0\), which are the electromagnetic waves, we will ignore here since they have nothing to do with our source.

### 4.3.1 Fourier transformations

We will determine the solution of the equation of motion by taking a Fourier transformation. Generally, we have

\[ f(x) = \int \frac{d^3p}{(2\pi)^3} \tilde{f}(p) e^{ix\cdot p} \]  

(4.86)

Recall, that the Fourier transformation of a \(\delta\)-function (or \(\delta\) tempered-distribution) is just

\[ \tilde{\delta}^3(p) = \int d^3x \delta^3(x) e^{-ix\cdot p} = 1 \]  

(4.87)

and

\[ \delta^3(x) = \int \frac{d^3p}{(2\pi)^3} e^{ix\cdot p} \]  

(4.88)

We obtain for the derivatives of \(\delta\)-functions

\[ \Delta \tilde{\delta}^3(x) = \int \frac{d^3p}{(2\pi)^3} \Delta e^{ix\cdot p} = \int \frac{d^3p}{(2\pi)^3} (ip)^2 e^{ix\cdot p} \]  

(4.89)

This also works for four-vectors

\[ \tilde{\delta}^4(p) = \int \frac{d^4k}{(2\pi)^4} e^{ik\cdot x} k_\mu \]  

(4.90)

and so

\[ \Box \tilde{\delta}^4(x) = \int \frac{d^4k}{(2\pi)^4} (-k_\mu k_\nu) e^{ix\cdot p} \]  

(4.91)

We will use this all the time: for you as a field theorist, \(\Box\) means \(-k^2\).
4.3.2 The Coulomb Potential

Back to our problem: since $\delta^3(\mathbf{x})$ is time-independent, we simplify

$$A_0(x) = \frac{e}{\Box} \delta^3(x) = -\frac{e}{\Delta} \delta^3(x) \quad (4.92)$$

We solve this using a Fourier transformation

$$A_0(x) = \int \frac{d^3p}{(2\pi)^3} \left( -\frac{e}{\Delta} \right) e^{ip \cdot x}$$

$$= e \int \frac{d^3p}{(2\pi)^3} \frac{1}{p^2} e^{ip \cdot x}$$

$$= \frac{e}{(2\pi)^2} \int_0^\infty dp \int_{-1}^1 d\cos \theta \int_0^{2\pi} d\phi \frac{1}{p} e^{ipr \cos \theta}$$

$$= \frac{e}{8\pi^2 \ir} \int_{-\infty}^{\infty} dp \frac{e^{-ipr} - e^{-ipr}}{p}$$

Note that the integrand does not diverge at $p \to 0$ since the numerator is $\propto \sin(pr) \approx pr + \ldots$. So we can introduce a small shift in the denominator $\ir$ and simplify

$$\int_0^{\infty} dp \frac{e^{ipr} - e^{-ipr}}{p} = \lim_{\epsilon \to 0} \left( \int_0^{\infty} dp \frac{e^{ipr} - e^{-ipr}}{p + \ir} \right)$$

Assuming $\epsilon > 0$, we see that the pole is on the negative imaginary axis. We will now use the residue theorem to calculate the integral: for $e^{ipr}$ we need to close the contour in the upper half-plane to get exponential suppression of the integrand $\sim e^{-|pr|}$, which misses the pole. For $e^{-ipr}$, we close the contour in the lower half plane and obtain

$$\int_{-\infty}^{\infty} dp \frac{-e^{-ipr}}{p + \ir} = (-2\pi \ir) \text{Res}_{p=-\ir} \left[ \frac{-e^{-ipr}}{p + \ir} \right] = (2\pi \ir) e^{-\epsilon r} \quad (4.93)$$

We therefore find

$$A_0(x) = \frac{e}{8\pi^2 \ir} (2\pi \ir) e^{-\epsilon r} = \frac{e}{4\pi \epsilon} e^{-\epsilon r} \quad (4.94)$$

and so for $\epsilon \to 0$, we obtain finally

$$A_0(x) = \frac{e}{4\pi r} \quad (4.95)$$

which is Coulomb's law.

4.3.3 A non-linear example: the graviton Lagrangian

Propagators are a useful tool to solve similar types of equations. Let us assume that our Lagrangian has a non-linear term $A^3$ (or as we will see: an interaction). Electro-magnetism does not have these terms and the equations of motion are linear, but there are many examples in nature, like the gluon or electro-weak theory. Another is the graviton. A heuristic Lagrangian for the graviton $h$ which is the fluctuation around a flat background metric $g_{\mu\nu}(x) = \eta_{\mu\nu} + h_{\mu\nu}(x) + \ldots$ can be written as

See Ex!
\[ L = -\frac{1}{2} h \Box h + \frac{1}{3} \lambda h^3 + J \cdot h \]  
(4.96)

where we have dropped the indices. This defines the non-linear equations of motion. We solve perturbatively and we start with the non-interacting theory with \( \lambda = 0 \)

\[ h_0 = \frac{1}{\Box} J \]  
(4.97)

We obtain the first order correction \( h = h_0 + h_1 \) where \( h_1 = \mathcal{O}(\lambda) \)

\[ \Box (h_0 + h_1) - \lambda (h_0 + h_1)^2 - J = 0 \]  
(4.98)

and using the 0th order result above,

\[ \Box h_1 = \lambda h_0^2 + \mathcal{O}(\lambda^2) \]  
(4.99)

This means that

\[ h_1 = \lambda \frac{1}{\Box} h_0^2 = \lambda \frac{1}{\Box} \left[ \left( \frac{1}{\Box} J \right) \left( \frac{1}{\Box} J \right) \right] \]  
(4.100)

In summary, the first order solution is

\[ h = h_0 + h_1 = \frac{1}{\Box} J + \lambda \frac{1}{\Box} \left[ \left( \frac{1}{\Box} J \right) \left( \frac{1}{\Box} J \right) \right] + \mathcal{O}(\lambda^2) \]  
(4.101)

This is the so-called **Green’s function method**. We will call the

\[ \Pi = -\frac{1}{\Box} \]  
(4.102)

the 2-point Green’s function or the **propagator**. If we define \( \Box_x = \partial^2 \) with respect to \( x^\mu \) and

\[ \Box_x \Pi(x, y) = -\delta^4(x - y) \]  
(4.103)

Note, \( \Pi(x, y) = \Pi(y, x) \). We can use Fourier transformations to show

\[ \Pi(x, y) = \int \frac{d^4 p}{(2\pi)^4} \frac{1}{p^2} e^{i(x-y) \cdot p} \]  
(4.104)

which we can check

\[ \Box_x \Pi(x, y) = - \int \frac{d^4 p}{(2\pi)^4} e^{i(x-y) \cdot p} = -\delta^4(x - y) \]  
(4.105)

We ignore subtleties with boundary conditions here but we will come back to this later on.

We can write the field as

\[ h(x) = \int d^4 y \, \delta^4(x - y) h(y) = - \int d^4 z \, (\Box_y \Pi(x, y)) h(y) \]  
(4.106)

\[ = - \int d^4 y \, \Pi(x, y) \Box_y h(y) \]  
(4.107)

where we have used integration by parts. We can solve the free field with \( \lambda = 0 \) equation, by inserting \( \Box_y h_0(y) = J(y) \) to obtain

\[ h_0(x) = - \int d^4 y \, \Pi(x, y) J(y) \]  
(4.108)
and the next order as
\[ \Box_w h_1(w) = \lambda h_0(w)^2 \]
\[ = \lambda \int d^4 y \Pi(w, y) J(y) \int d^4 z \Pi(w, z) J(z) \]
(4.109)
(4.110)

Substituting again as above in Eq. (4.107) we find for the perturbative solution
\[ h(x) = -\int d^4 y \Pi(x, y) J(y) - \lambda \int d^4 s \int d^4 y \int d^4 y \Pi(x, s) \Pi(s, y) \Pi(s, z) J(y) J(z) + O(\lambda^2) \]
which is the meaning of Eq. (4.101).

There is a nice pictorial representation of this solution in the form of Feynman diagrams. This pictorial representation allows us to
\[ h(x) = \Pi(x, y) J(y) + h(x) \]
\[ + \ldots \]

derive the most general \( \lambda^n \) expansion for a classical field with the following classical Feynman rules:

1. Draw a line from a point \( x \) to a new point \( x_j \).
2. Truncate with a source \( J(x_i) \) or branch into two lines and multiply with a factor \( \lambda \).
3. Repeat the previous step.
4. The final result for \( h(x) \) is given by all graphs up to some order \( \lambda^n \) with the ends truncated by currents \( J(x_i) \) and the lines replaced by propagators \( \Pi(x_i, x_j) \). All internal points are integrated over.

As we will see, in quantum field theory the Feynman rules are almost identical except for \( \hbar \to 0 \) (classical limit), lines can not close in on themselves.\(^{17}\) In classical general relativity, these diagrams describe the post-Newtonian effects of the Sun on Mercury! The lines represent gravitons and the sources are in this case the Sun. The first diagram is the well-known Newtonian potential while the second order represents the self-interaction \( \lambda \sim \sqrt{G_N} \) which is a prediction of general relativity.

**4.4 The Hamiltonian Formalism**

In this course we will not discuss path integrals, and focus instead on canonical quantization. For this we need the Hamiltonian formalism of field theory, since canonical commutators are defined between the field and its canonically conjugate momentum.

We first define the conjugate momentum \( \pi^a(x) \)
\[ \pi^a(x) = \frac{\partial L}{\partial \dot{\phi}_a} \]
(4.111)
The Hamilton density is obtained by a Legendre transformation
\[ H = \pi^a(x)\dot{\phi}_a(x) - \mathcal{L}(x) \quad (4.112) \]
As in classical mechanics, we replace \( \dot{\phi}_a(x) \) in favor of \( \pi^a(x) \) everywhere. The Hamiltonian is simply
\[ H = \int d^3x \mathcal{H} \quad (4.113) \]

### 4.4.1 Example: a real scalar field

For the Lagrangian
\[ \mathcal{L} = \frac{1}{2} (\partial_t \phi)^2 - \frac{1}{2} (\nabla \phi)^2 - \frac{1}{2} m^2 \phi^2 \]
we obtain \( \pi = \dot{\phi} \) and the Hamiltonian
\[ \mathcal{H} = \frac{1}{2} \pi^2 + \frac{1}{2} (\nabla \phi)^2 + \frac{1}{2} m^2 \phi^2 \]
The advantage of the Lagrangian formalism is its explicit Lorentz invariance. In contrast, the Hamiltonian formalism is not manifestly Lorentz invariant, since we have picked a preferred time. The equations of motions do not look Lorentz invariant
\[ \partial_t \phi(x, t) = \frac{\partial \mathcal{H}}{\partial \pi(x, t)} \quad , \quad \partial_t \pi(x, t) = -\frac{\partial \mathcal{H}}{\partial \phi(x, t)} \]
The physics is of course unchanged (and still Lorentz invariant). In the canonical formalism, we will have to check at various points that that the invariance is indeed preserved.
5

Canonical Quantization

5.1 Free fields

You are very familiar with the transition from classical physics to quantum mechanics by using canonical quantization: it is a recipe based on the Hamiltonian formalism. We take the generalized coordinates and their conjugate momenta \( q_a, p^a \) and promote them to hermitian operators. This recipe is closest to classical mechanics, when we formulate the dynamics using Poisson brackets \( \{ q_b, p^a \} = \delta^a_b \) and morph them into commutation relations between operators

\[
[q_a, p^b] = i \hbar \delta^a_b \\
[q_a, q_b] = 0 \\
[p^a, q^b] = 0
\]

As in the introductory example of the discrete chain, we will do the same now for the field \( \phi_a(x) \) and its momentum conjugate \( \pi^b(x) \).

A quantum field operator valued function of space obeys the commutation relations

\[
[\phi_a(x), \pi^b(y)] = i \hbar \delta^a_b \delta^3(x - y) \\
[\phi_a(x), \phi_b(y)] = 0 \\
[\pi^a(x), \pi^b(y)] = 0
\]

We are working in the Schrödinger picture so that the operators \( \phi_a(x), \pi^b(x) \) do not depend on time at all. Note, that we have lost track of Lorentz invariance since we have separated space and time. All time dependence in the Schrödinger picture is encapsulated in the states \( |\psi\rangle \) which evolve according to the Schrödinger equation

\[
i \hbar \frac{d}{dt} |\psi\rangle = H |\psi\rangle
\]

This is exactly the same as in quantum mechanics, we are just applying the formalism to the fields. Warning: the notation \( |\psi\rangle \) is deceiving since the wave function in quantum field theory is a functional of every possible configuration of the field \( \phi_a \).

In quantum mechanics we would determine the spectrum of the Hamiltonian \( H \) and obtain the time evolution this way. This is too

This is also often confusingly referred to as second quantization. However, the fact that there are discrete modes is a classical phenomenon (think of the eigenmodes of a string). The two steps are in fact (1) interpret these modes as having energy \( E = \hbar \omega \) and (2) quantize each mode as a harmonic oscillator. In that sense we are only quantizing once.
hard in quantum field theories (most of the time). Clearly, dealing with an infinite number of degrees of freedom complicates the problem somewhat. However for free theories, we can solve the dynamics by going to a basis where we can separate the dynamics of the degrees of freedom. Typically, these will be Lagrangians which are quadratic in the fields.

The simplest relativistic field theory is the Klein-Gordon equation for a real scalar field \( \phi(x, t) \)

\[
\partial_\mu \partial^\mu \phi + m^2 \phi = 0
\]

To decouple the degrees of freedom, we only need to take the Fourier transformation

\[
\phi(x, t) = \int \frac{d^3p}{(2\pi)^3} \phi(p, t) e^{i x \cdot p}
\]

Then

\[
\left[ \frac{\partial^2}{\partial t^2} + (p^2 + m^2) \right] \phi(p, t) = 0 \tag{5.6}
\]

For each \( p \), \( \phi(p, t) \) solves the harmonic oscillator equation with a frequency

\[
\omega_p = +\sqrt{p^2 + m^2}
\]

The most general solution is a linear superposition of simple harmonic oscillators, each vibrating at a different frequency with a different amplitude. To quantize this theory we must quantize these infinitely many harmonic oscillators. You should read Section 2.2.3 again.

### 5.1.1 The free scalar field

We write the field and its conjugate momentum as a linear sum of infinitely many creation \( a^\dagger_p \) and annihilation \( a_p \) operators labeled by their 3-momentum \( p \)

\[
\phi(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_p}} \left[ a_p e^{ix \cdot p} + a^\dagger_p e^{-ix \cdot p} \right] \tag{5.9}
\]

and for the conjugate momentum\(^3\)

\[
\pi(x) = \int \frac{d^3p}{(2\pi)^3} (-i) \frac{1}{\sqrt{2\omega_p}} \left[ a_p e^{ix \cdot p} - a^\dagger_p e^{-ix \cdot p} \right] \tag{5.10}
\]

As in the case of the quantized string, the commutation relations for \( \phi(x) \) and \( \pi(x) \) of Eq. (5.1) are equivalent to commutation relations for \( a_p \) and \( a^\dagger_p \). Quantum field theory is just quantum mechanics with an infinite number of harmonic oscillators.

\[
[a_p, a^\dagger_q] = (2\pi)^3 \delta^{(3)}(p - q) \tag{5.11}
\]

\[
[a_p, a_q] = [a^\dagger_p, a^\dagger_q] = 0 \tag{5.12}
\]

---

1 At least one for each point \( x \) in space.

2 Recall the eigen-modes of the discrete chain.

3 Inspired by Eq. (2.51) or Eq. (5.8).

You can check that this will lead to the correct commutation relations for the fields consistent with the ladder operators.

Later, when we will consider time-dependent field operators (Heisenberg picture), you will see that this is consistent with

\[
\pi = \frac{\partial L}{\partial (\partial_\phi)} = \partial_t \phi
\]
We will show this one way

\[
[\phi(x), \pi(y)] = \int \frac{d^3 p}{(2\pi)^3} \int \frac{d^3 q}{(2\pi)^3} - i \frac{\omega_q}{\omega_p} \left[ - \frac{[a_p, a_q^\dagger]}{(2\pi)^3} e^{ixp-iyq} + \frac{[a_p^\dagger, a_q]}{(2\pi)^3} e^{-ixp+iyq} \right]
\]

\[
= \int \frac{d^3 p}{(2\pi)^3} - i \frac{\omega_q}{\omega_p} \left[ - e^{i(x-y)p} - e^{-i(x-y)p} \right]
\]

\[
= i \int \frac{d^3 p}{(2\pi)^3} e^{i(x-y)p}
\]

\[
= i \delta^{(3)}(x-y)
\]

We obtain for the Hamiltonian in terms of \(a_p, a_q^\dagger\)

\[
H = \frac{1}{2} \int d^3 x \left[ \pi(x)^2 + (\nabla \phi(x))^2 + \frac{m^2}{2} \phi(x)^2 \right]
\]

\[
= \frac{1}{2} \int d^3 x \frac{d^3 p}{(2\pi)^3} \frac{d^3 q}{(2\pi)^3} (1) \sqrt{\frac{\omega_p \omega_q}{\omega_p \omega_q}} \left[ a_p e^{ixp} - a_p^\dagger e^{-ixp} \right] \left[ a_q e^{iyq} - a_q^\dagger e^{-iyq} \right]
\]

\[
+ \frac{1}{\sqrt{\omega_p \omega_q}} \left[ ip a_p e^{ixp} - ip a_p^\dagger e^{-ixp} \right] \left[ iq a_q e^{iyq} - iq a_q^\dagger e^{-iyq} \right]
\]

\[
+ m^2 \frac{1}{\sqrt{\omega_p \omega_q}} \left[ a_p e^{ixp} + a_p^\dagger e^{-ixp} \right] \left[ a_q e^{iyq} + a_q^\dagger e^{-iyq} \right]
\]

We will now collect the terms, integrate over \(d^3 x\) and replace \(e^{-ix(p+p)} \rightarrow \delta^3(p-p)\) and \(e^{-ix(p-p)} \rightarrow \delta^3(p+p)\) which allows us to remove one \(d^3 q\) integral.

\[
H = \int \frac{d^3 p}{(2\pi)^3} \frac{1}{4\omega_p} \left[ a_p a_{-p} + a_p^\dagger a_{-p}^\dagger \right] (-\omega_p^2 + p^2 + m^2)
\]

\[
+ \frac{1}{4\omega_p} \left[ a_p a_p^\dagger + a_p^\dagger a_p \right] (\omega_p^2 + p^2 + m^2)
\]

The first line vanishes because of \(\omega_p^2 = p^2 + m^2\) and we obtain

\[
H = \frac{1}{2} \int \frac{d^3 p}{(2\pi)^3} \omega_p \left[ a_p a_p^\dagger + a_p^\dagger a_p \right]
\]

\[
= \int \frac{d^3 p}{(2\pi)^3} \omega_p \left[ a_p a_p^\dagger + \frac{1}{2} (2\pi)^3 \delta^{(3)}(0) \right]
\]

What is this? We clearly have to work to interpret this properly. There are two endings: the delta-function has an infinite spike at 0 and the integral over \(\omega_p\) also diverges.\(^4\)

\[5.1.2 \textbf{Infinites} \]

We will separate the infinites into two classes: **infrared divergences** which are caused by assuming infinitely large scales (or arbitrarily low energies) and **ultraviolet divergences** which come from extrapolating to infinitely small scales (or arbitrarily high energies).

We will better understand the problem by discussing the ground state or the vacuum \(|0\rangle\). We define \(|0\rangle\) as the state being annihilated by all the \(a_p\)

\[
a_p |0\rangle = 0 \quad \text{for all } p
\]

\[5.15\]
We obtain for the ground-state energy
\[ H|0\rangle = E_0|0\rangle = \int \frac{d^3p}{(2\pi)^3} \omega_p \left[a_p^\dagger a_p + \frac{1}{2} (2\pi)^3 \delta^{(3)}(0)\right]|0\rangle \]  
(5.16)
\[ = \int \frac{d^3p}{(2\pi)^3} \omega_p \frac{1}{2} (2\pi)^3 \delta^{(3)}(0)|0\rangle \]  
(5.17)
\[ = \cdot \cdot \cdot |0\rangle \]  
(5.18)

This is common for quantum field theories: infinities pop up left and right. They tell us that we might be asking the wrong question or are doing something wrong.

**The infrared divergence** or long-distance divergence can be taken care of by putting the field in a box of length \( L \) or volume \( V = L^3 \) and assuming periodic boundary conditions. We obtain
\[ (2\pi)^3 \delta^{(3)}(0) = \lim_{L \to \infty} \int_{-\frac{L}{2}}^{\frac{L}{2}} dx \int_{-\frac{L}{2}}^{\frac{L}{2}} dy \int_{-\frac{L}{2}}^{\frac{L}{2}} dz e^{ixp}|p=0 = V \]  
(5.19)
This tells us: we should be calculating the energy density \( \mathcal{E}_0 \) instead of the total energy! We therefore divide by \( V \) to find
\[ \mathcal{E}_0 = \frac{E}{V} = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2} \omega_p = \frac{4\pi}{(2\pi)^3} \frac{1}{2} \int_0^\infty dp \frac{1}{\sqrt{m^2 + p^2}} \to \infty \]  
(5.20)
which is however still infinite.\(^5\)

**The ultraviolet divergence** or short-distance divergence can be understood by recognizing this as the integral over all ground state energies of the harmonic oscillators. In the limit that \( |p| \to \infty \) we find \( \mathcal{E}_0 \to \infty \). This divergence is caused by us assuming that our quantum field theory is valid down to the smallest distances (or highest energies)\(^6\). This is certainly not true. We need to cut-off the integral at high energy (or high momentum) to include the fact that we are ignorant above a certain scale.

We could introduce an *explicit cut-off* and stop integrating at some energy \( \Lambda \) since we do not want to extrapolate beyond this point.
\[ \mathcal{E}_0(\Lambda) = \frac{1}{4\pi^2} \int_0^\Lambda dp \frac{p^2}{\sqrt{m^2 + p^2}} \propto a\Lambda^4 + \ldots \]  
(5.21)
which is called *regularization*.\(^7\)

We can then *renormalize* the vacuum energy to its observed value by realizing that we could have added a constant \( V_0 \) to the Lagrangian
\[ \mathcal{L}_{\text{new}} = \mathcal{L} - V_0 \]  
(5.22)
This constant will not affect the equations of motion and we usually just drop it. \( V_0 \) is a *bare* parameter which we choose to cancel the divergence of zero-point energies
\[ V_0 = -\mathcal{E}_0(\Lambda) + \chi \]  
(5.23)

\(^5\) Although a little less infinite: "\( \infty \) "...  

\(^6\) Even above the Planck scale! Compare to the discrete chain, where this would correspond to extrapolating to distances much smaller than the average distance between the mass points

\(^7\) Note, that this is not a Lorentz-invariant regulator: Lorentz invariance is broken when we impose a cut-off on the spatial momentum only. To meaningfully evaluate the zero-point energy density, you must use a regularization scheme that does not break Lorentz invariance. A possible choice is dimensional regularization which will be covered in more advanced QFT courses.
this diverges as we take the cut-off to infinity \( V_0 = -a\Lambda^4 + \ldots \). But now the **observed** energy-density is finite even for \( \Lambda \to \infty \) as

\[
\frac{1}{V} H \, |0\rangle = (E_0(\Lambda) + V_0)|0\rangle = (E_0(\Lambda) - E_0(\Lambda) + \chi)|0\rangle = E_{\text{observed}} |0\rangle
\]

the step is called **renormalisation**. Note that the finite piece \( \chi = E_{\text{observed}} \) is completely arbitrary. It must be determined experimentally by measuring an observable – in this case the vacuum energy.\(^8\) Since \( \Lambda \) only appears in the "bare" Lagrangian and intermediate steps of the calculation, but not in physical observables, we can remove it at the end.

### 5.1.3 Normal ordering

Alternatively, we can follow a more practical approach called **normal ordering**. Since we are only interested in energy-differences\(^9\), the practical approach is good enough. We will just subtract off the infinity and redefine our Hamiltonian as

\[
H' = \int \frac{d^3p}{(2\pi)^3} \omega_p a_p^\dagger a_p
\]

and now

\[
H'|0\rangle = 0
\]

If you are a bit taken aback by the lack of rigor implied by an arbitrary redefinition of our Hamilton, note the following: there is an inherent **ambiguity of ordering** transitioning from classical to quantum. If we had defined the classical Hamiltonian as

\[
H'_{\text{classical}} = \frac{1}{2} (\omega q - ip)(\omega q + ip) = \frac{1}{2} p^2 + \frac{\omega}{2} q^2 = H_{\text{classical}}
\]

which is classically equivalent \( H'_{\text{classical}} = H_{\text{classical}} \). We would naturally get after quantization\(^10\)

\[
H'_{QM} = \omega a^\dagger a
\]

This method of exploiting the order ambiguity is called **normal ordering**: we define it as the ordering in which all the creation operators stand to the left of the annihilation operators.

\[
(a_p a_{q\dagger})_{\text{normal ordered}} = :a_p a_{q\dagger}: = a_{q\dagger}^\dagger a_p
\]

\[
(a_{p\dagger} a_q)_{\text{normal ordered}} = :a_{p\dagger} a_q: = a_{p\dagger} a_q
\]

In general, we define a string of operators as normal ordered

\[
:\phi_a(x_1) \phi_b(x_2) \ldots \phi_f(x_n) :
\]
The Cosmological Constant: As mentioned before, gravity can indeed measure energy differences. Gravity sees everything! We would expect the sum of all zero-point energies to contribute to the cosmological constant \( \Lambda = E_0/V \) and appear on the right side of the Einstein-Hilbert equations

\[
R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = -8\pi G T_{\mu\nu} + \Lambda g_{\mu\nu}.
\]

Cosmological observations tell us that ca. 70% of the energy density of the universe today is in a term that looks like the cosmological constant with

\[
\Lambda \sim (10^{-3} \text{eV})^4
\]

This is many orders of magnitude smaller than all the other scales in particle physics (the SM is valid to \( \sim 200 \text{ GeV} = 2 \cdot 10^{11} \text{eV} \)). Why don’t zero-point energies contribute to \( \Lambda \)? What cancels them to such accuracy? We do not know the answer. In fact, we don’t even know how to ask the right question.

if all the annihilation operators are placed to the right. Another consequence of normal-ordering is that vacuum matrix elements vanish

\[
\langle 0 | : \phi_a(x_1) \phi_b(x_2) \ldots \phi_f(x_n) : | 0 \rangle
\]

and the only normal-ordered expressions that do not vanish are c-number functions\(^{11}\) which satisfy

\[
\langle 0 | : f(x) : | 0 \rangle = f(x)
\]

Our Hamiltonian then has the normal-ordered form

\[
: H : = \int \frac{d^3 p}{(2\pi)^3} \omega_p a_p^\dagger a_p
\]  

(5.30)

5.2 Particles

What are the excitations of the field with respect to the vacuum \( |0\rangle \)? We know that

\[
[H, a_p^\dagger] = \omega_p a_p^\dagger
\]

\[
[H, a_p] = -\omega_p a_p
\]

As in the case of the simple harmonic oscillator, we can construct energy eigenstates by applying the creation operators \( a_p^\dagger \) on the vacuum \( |0\rangle \). We define

\[
a_p^\dagger |0\rangle = \frac{1}{\sqrt{2E_p}} |p\rangle
\]  

(5.33)

The factor \( \sqrt{2E_p} \) constitutes a relativistic normalization – see below.\(^{12}\) The energy of this state is

\[
: H : |q\rangle = \int \frac{d^3 p}{(2\pi)^3} \omega_p a_p^\dagger a_p \sqrt{2E_q} a_q^\dagger |0\rangle
\]

\[
= \sqrt{2E_q} \int \frac{d^3 p}{(2\pi)^3} \omega_p a_p^\dagger a_p \left( a_q^\dagger a_p + (2\pi)^3 \delta^{(3)}(p - q) \right) |0\rangle
\]

\[
= \omega_q \sqrt{2E_q} a_q^\dagger |0\rangle
\]

\[
= \omega_q |q\rangle
\]

Ex: Verify the relation

\[
[H, a_p^\dagger] = \omega_p a_p^\dagger
\]

11 The term c-number (or classical number) is an old nomenclature used by Paul Dirac which refers to real and complex numbers. It is used to distinguish from operators (q-numbers or quantum numbers) in quantum mechanics.

12 The \( \sqrt{2} \) is for convenience only.
This is the relativistic dispersion relation (or relativistic Pythagoras) for a particle of mass \( m \) and momentum \( \mathbf{q} \). We therefore interpret this as a momentum eigenstate of a single particle of mass \( m \). Its energy is \( E^2_q = \mathbf{q}^2 + m^2 \) and we can interchange \( E_q \) and \( \omega_q \).

What about the other quantum numbers? The total momentum \( \mathbf{P} \) of a classical field configuration which we have derived in Eq. (4.73), can be turned into a (normal ordered) operator:

\[
\mathbf{P} : = - \int \frac{d^3 p}{(2\pi)^3} \frac{d^3 q}{(2\pi)^3} (i) \sqrt{\frac{\omega_p}{4\omega_q}} \left[ a_p e^{ix \cdot \mathbf{p}} - a^\dagger_p e^{-ix \cdot \mathbf{p}} \right] \nabla \left[ a_q e^{ix \cdot \mathbf{q}} + a^\dagger_q e^{-ix \cdot \mathbf{q}} \right] : 
\]

The first term vanishes because the term in the brackets is symmetric under \( \mathbf{p} \to -\mathbf{p} \) which together with \( \mathbf{p} \) makes this part of the integrand anti-symmetric and therefore vanishes after integrating over a symmetric interval. We find

\[
\mathbf{P} : = \int \frac{d^3 p}{(2\pi)^3} \mathbf{p} a^\dagger_p a_p
\]

In summary we find

\[
\mathbf{P}^\mu : = \int \frac{d^3 p}{(2\pi)^3} p^\mu a^\dagger_p a_p
\]

We can do the same for the classical expression of the angular momentum\(^ {13} \) to obtain the operator

\[
\mathbf{J}^i = \varepsilon^{ijk} \int d^3x (\mathbf{J}^0)^jk
\]

Acting on a one-particle state with zero-momentum, we find

\[
\mathbf{J}^i |\mathbf{p} = 0\rangle = 0
\]

which we understand as telling us that the particle does not carry internal angular momentum. Quantizing a scalar field gives rise to a spin-0 particle!

\[5.2.1\] Multi-particle states and Bose statistics

Similarly, we can see that the state

\[
a^\dagger_{p_1} a^\dagger_{p_2} \ldots a^\dagger_{p_n} |0\rangle
\]
has momentum
\[ \mathbf{p}_1 + \mathbf{p}_2 + \ldots + \mathbf{p}_n \] (5.39)

We interpret this as an \textit{n-particle state}.
\[ |\mathbf{p}_1, \ldots, \mathbf{p}_n\rangle = \mathcal{N} a^\dagger_{\mathbf{p}_1} a^\dagger_{\mathbf{p}_2} \ldots a^\dagger_{\mathbf{p}_n} |0\rangle \] (5.40)
where \( \mathcal{N} \) is a normalization. Since the creation operators commute \([a^\dagger_{\mathbf{p}_1}, a^\dagger_{\mathbf{p}_2}] = 0\), these states are \textit{symmetric under exchange} of any two particles, e.g.
\[ |\mathbf{p}_1, \mathbf{p}_2\rangle = |\mathbf{p}_2, \mathbf{p}_1\rangle \] (5.41)
and we know that these particles are \textbf{bosons}. We can construct the \textbf{Fock space} as the Hilbert space spanned by all possible combinations of the \( a^\dagger_{\mathbf{p}} \),
\[ |0\rangle, a^\dagger_{\mathbf{p}_1}|0\rangle, a^\dagger_{\mathbf{p}_1}a^\dagger_{\mathbf{p}_2}|0\rangle, a^\dagger_{\mathbf{p}_1}a^\dagger_{\mathbf{p}_2}a^\dagger_{\mathbf{p}_3}|0\rangle, \ldots \] (5.42)
As we have seen in the introduction, this is simply the direct sum of the \( n \)-particle Hilbert spaces
\[ |\Psi\rangle = |\Psi\rangle_1 \oplus |\Psi\rangle_2 \oplus \ldots \] (5.43)
We can count the number of particles in a given state in Fock space using the \textbf{number operator} \( N \) defined as
\[ N = \int \frac{d^3p}{(2\pi)^3} a^\dagger_{\mathbf{p}} a_{\mathbf{p}} \] (5.44)
which satisfies
\[ N |\mathbf{p}_1, \ldots, \mathbf{p}_n\rangle = n |\mathbf{p}_1, \ldots, \mathbf{p}_n\rangle \] (5.45)
and it commutes with Hamiltonian,
\[ [H, N] = 0 \]
which it should since in our free theory, particle number is conserved. This will no longer hold once we include interactions (e.g. \( \mathcal{L} \sim \phi^3 \)). Interactions can destroy and annihilate particles, taking us between different sectors in Fock space.

5.3 \textbf{Smeared operator valued distributions}
As already mentioned, \( |\mathbf{p}\rangle \) is not really a particle and not localized in space but \( |\mathbf{p}\rangle \) is a momentum eigenstate. Crucially, neither the operators \( \phi(x) \) nor \( a^\dagger_{\mathbf{p}} \) are good operators acting on the Fock space. \footnote{This is in direct analogy with the position and momentum eigenstates of quantum mechanics which are not good elements of the Hilbert space because they are not normalizable. They normalize to delta-functions.}
\[ \langle \mathbf{p}|\mathbf{p}\rangle = 2E_p \langle 0|a^\dagger_{\mathbf{p}}a^\dagger_{\mathbf{p}}|0\rangle \]
\[ = 2E_p \langle 0|a^\dagger_{\mathbf{p}}a_{\mathbf{p}} + (2\pi)^3\delta^{(3)}(\mathbf{p} - \mathbf{p})|0\rangle \]
\[ = (2\pi)^32E_p \delta^{(3)}(\mathbf{0}) \langle 0|0\rangle \]
\[ = 1 \]
and so

\[ \langle \mathbf{p}\mathbf{p} \rangle = 2E_\mathbf{p} (2\pi)^3 \delta^{(3)}(0) \]  

(5.46)

Similarly

\[ \langle 0|\phi(x)\phi(x)|0 \rangle = \langle x|x \rangle = \delta^{(3)}(0) \]  

(5.47)

These are therefore **operator valued distributions**, rather than functions. Even though we obtain a finite vacuum expectation value

\[ \langle 0|\phi(x)|0 \rangle = 0 \]  

(5.48)

The quadratic fluctuations of the operator at a point are infinite

\[ \langle 0|\phi(x)\phi(x)|0 \rangle = \infty \]

We can obtain finite results by smearing the distributions over space (or momentum) using test functions,

\[ \phi_f(x) = \int d^3y \phi(x)f(x - y) \]

with e.g. the (normalized) Gaussian smearing

\[ f(x) = \frac{1}{(a^2\pi)^{3/2}} \exp \left( -\frac{x^2}{a^2} \right) \]

We can obtain \( \langle 0|\phi_f(x)\phi_f(x)|0 \rangle \) by observing that of the four terms after expanding \( \phi(x)^2 \) only the one proportional to \( \langle 0|a_p a_q^\dagger|0 \rangle \sim \delta^3(p - q) \) will be non-zero. We have

\[ \langle 0|\phi_f(x)\phi_f(x)|0 \rangle = \frac{1}{8(a\pi)^3} \int d^3p \left[ \int d^3y \exp \left( -\frac{(x - y)^2}{a^2} + ip(x - y) \right) \right]^2 \]

The Gaussian integral can be easily calculated by completing the squares\(^{15}\) and we get

\[ \langle 0|\phi_f(x)\phi_f(x)|0 \rangle = \frac{1}{(2\pi)^3} \int \frac{d^3p}{2E_\mathbf{p}} e^{-p^2a^2/2} \]

\[ = \frac{1}{(2\pi)^2} \int_0^\infty dp \frac{p^2}{\sqrt{\mu^2 + m^2}} e^{-p^2a^2/2} \]

For \( a \ll 1/m \) we can ignore the \( m \)-dependence\(^{16}\) to obtain

\[ \langle 0|\phi_f(x)\phi_f(x)|0 \rangle = \frac{1}{4\pi^2} \frac{1}{a^2} \]

We see that the fluctuations diverge without bounds in the limit of \( a \to 0 \).

5.3.1 **Dimensional analysis and Coleman theorem**

We could have in fact guessed the result for the fluctuations easily. In the limit \( a \to 0 \), there is no scale\(^{17}\) other than \( a \). Similarly, we know by counting the dimensions in the expansion of \( \langle 0|\phi(x)\phi(y)|0 \rangle \) and realizing again that the only scale is \( |x - y| \), we have to have

\[ \langle 0|\phi(x)\phi(y)|0 \rangle \sim \frac{1}{|x - y|^2} \]

Recall, \( a_p(0) = 0 \) and therefore by hermitian conjugation \( \langle 0|a_p^\dagger|0 \rangle = 0 \).
This **dimensional analysis** is equivalent to saying that the field has dimensions of inverse length. We could have already realized that a scalar field in \(3 + 1\) Dimensions has canonical mass dimension\(^1\), or
\[
\frac{1}{18}
\]
The mass dimension of a field \(\psi\) is defined to be the exponent
\[
[\psi] = n \sim (\text{mass})^n
\]
which means that \(\psi\) has dimension \((\text{mass})^n \sim (\text{energy})^n \sim (\text{length})^{-n}\).

Since the action is dimensionless \([S] = [\hbar] = 0\) and since \([d^4x] = -4\) and \([\partial_\mu] = 1\), we find
\[
[\phi(x)] = 1
\]
So the growth of the fluctuations with \(|x - y|^{-2}\) is really just dimensional analysis. We can obtain similar results in any dimension. In \(D\) space-time dimensions
\[
S = \int d^Dx \partial_\mu \phi(x) \partial^\mu \phi(x) + \ldots
\]
from which we get
\[
[\phi(x)] = \frac{D - 2}{2}
\]
and so in \(D\)–dimensions we obtain
\[
\langle 0 | \phi(x) \phi(y) | 0 \rangle_D \sim \frac{1}{|x - y|^{D - 2}}
\]
We immediately see an interesting feature by going to lower dimensions. The divergence of the fluctuations is less rapid. In particular for \(D = 2\), that is in \(1 + 1\) space-time dimensions, the power vanishes. This is really a log and we get\(^19\)
\[
\langle 0 | \phi(x) \phi(y) | 0 \rangle_{D=2} \sim \ln \left( \frac{1}{|x - y|} \right)
\]
The fluctuations diverge logarithmically slow as \(|x - y| \to 0\). The flip side of this is that also for **large** \(|x - y|\), the fluctuations diverge logarithmically.

So long-range fluctuations in \(D \leq 2\) are then actually very important now and if you follow this through you find that these fluctuations are large enough to prevent the formation of long-range order.\(^20\) This is a form of the **Coleman-Mermin-Wagner** theorem,\(^21\) which holds in quantum field theory and in statistical mechanics. They hold in both since we can get statistical mechanics from quantum field theory by analytic continuation – but this is for another lecture.

### 5.4 Some technicalities: relativistic normalization

The vacuum \(|0\rangle\) is defined as normalized \(\langle 0 | 0 \rangle = 1\). Our one-particle \(|p\rangle = \sqrt{2E_p}a_p\dagger |0\rangle\) states then have
\[
\langle p | q \rangle = (2\pi)^3 \sqrt{4E_pE_q} \delta^{(3)}(p - q)
\]

\(^{18}\) The **mass dimension** of a field \(\psi\) is defined to be the exponent
\[
[\psi] = n \sim (\text{mass})^n
\]

\(^{19}\) See Exercise 5.2.

\(^{20}\) which in particle physics language is called **spontaneous symmetry breaking**.

\(^{21}\) Continuous symmetries cannot be spontaneously broken in dimensions \(D \leq 2\), [https://en.wikipedia.org/wiki/Mermin-Wagner_theorem](https://en.wikipedia.org/wiki/Mermin-Wagner_theorem)
This is a Lorentz invariant normalization. How do we show this? In a quantum theory, we are always looking for **unitary representations** of our symmetries, such that

\[ p^\mu \rightarrow \tilde{p}^\mu = \Lambda^\mu_{\nu} p^\nu \quad |p\rangle \rightarrow |\tilde{p}\rangle = U(\Lambda)|p\rangle \]

and therefore leave the normalizations unchanged

\[ (p|p) = (\tilde{p}|\tilde{p}) \]

but since \( p \) is a three-vector, this is not at all clear. The problem is that spatial volumes are not invariant under boosts \( V \rightarrow V/\gamma \)!

We show that we have a Lorentz invariant normalization by proving that another object is Lorentz invariant.

**Claim:** We claim that

\[ \int \frac{d^3p}{2E_p} \quad (5.49) \]

is the Lorentz invariant measure.

**Proof:** We show this by using that \( \int d^4p \) is obviously Lorentz invariant and by using that

\[ p^\mu p_\mu = m^2 \quad \Rightarrow \quad p_0^2 = E_p^2 = p^2 + m^2 \]

is a Lorentz invariant statement. We choose the positive branch \( p_0 = +E_p \), but Lorentz transformations (the ones which are continuously connected to the \( \mathbb{1} \)) also leave this choice invariant. We conclude that the following combination must be Lorentz invariant

\[ \int d^4p \delta(p_0^2 - p^2 - m^2) \bigg|_{p_0 > 0} \]

which completes the proof. Now we know that the Lorentz-invariant \( \delta^{(3)} \)-function for 3 vectors is:

\[ 2E_p \delta^{(3)}(p - q) \]

this holds because

\[ \int \frac{d^3p}{2E_p} 2E_p \delta^{(3)}(p - q) = 1 \]

This finally explains our normalization for the momentum states,

\[ |p\rangle = \sqrt{2E_p} \alpha_p^\dagger |0\rangle \]

These states now satisfy

\[ \langle q|p\rangle = (2\pi)^3 2E_p \delta^{(3)}(p - q) \quad (5.50) \]

The identity on one-particle states is therefore

\[ \mathbb{1} = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_p} |p\rangle\langle p| \]

In the following, we will only use relativistically invariant normalizations.

Note, that

\[ \delta(g(x)) = \sum_i \frac{\delta(x - x_i)}{|g'(x_i)|} \]

where the sum extends over all roots of \( g(x) \) which are assumed to be simple. For example

\[ \delta(x^2 - \alpha^2) = \frac{1}{2|\alpha|} \left[ \delta(x + \alpha) + \delta(x - \alpha) \right]. \]
5.4.1 Meaning of the state $\phi(x)|0\rangle$

Let us consider the meaning of $\phi(x)|0\rangle$. Using the expansion Eq. (5.9) and Eq. (5.33), we obtain

$$\phi(x)|0\rangle = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_p} e^{-ip \cdot x} |p\rangle$$

which is a linear superposition of single particle states with well-defined momentum. This should remind you of the non-relativistic expression for the eigen-state of position $|x\rangle$, except for the factor of $1/2E_p$. We observe that for non-relativistic $p$ the factor is nearly a constant $1/2m$.

We will therefore put forward the same interpretation: the operator $\phi(x)$ acting on the vacuum creates a particle at position $x$.

Additional evidence comes from computing

$$\langle 0|\phi(x)|p\rangle = \langle 0| \int \frac{d^3p'}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \left[ a_{p'} e^{ix \cdot p'} + a^\dagger_{p'} e^{-ix \cdot p'} \right] \sqrt{2E_p} a^\dagger_p |0\rangle = e^{ix \cdot p}$$

This can be interpreted as the position-space representation of the single-particle wavefunction of the state $|p\rangle$, exactly like in non-relativistic quantum mechanics $\langle x|p\rangle \propto e^{ix \cdot p}$ is the wavefunction of the state $|p\rangle$.

5.5 Complex scalar fields

Let us consider a complex scalar field $\psi(x)$ with

$$\mathcal{L} = \partial_\mu \psi^* \partial^\mu \psi - m^2 \psi^* \psi$$

with the equation of motion

$$\partial_\mu \partial^\mu \psi + m^2 \psi = 0$$

and the same equation of motion for $\psi^*$. We expand the complex field operator in the Schrödinger picture

$$\psi(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \left[ b_p e^{ix \cdot p} + c^\dagger_p e^{-ix \cdot p} \right]$$

$$\psi^\dagger(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \left[ b^\dagger_p e^{-ix \cdot p} + c_p e^{ix \cdot p} \right]$$

The classical field is not real: the corresponding quantum field $\psi$ is not hermitian. This explains the appearance of the different operators $b$ and $c^\dagger$. The conjugate momentum is

$$\pi = \frac{\partial \mathcal{L}}{\partial \partial_t \psi} = \partial_t \psi^*$$

and\footnote{Recall, that we define $\pi(x)$ in analogy with the result for the harmonic oscillator and that it satisfies the canonical commutation relations.}
\[ \pi(x) = \int \frac{d^3p}{(2\pi)^3} \left( \frac{E_p}{2} \right)^{1/2} \left[ b_p^\dagger e^{-ix\cdot p} - c_p e^{ix\cdot p} \right] \] (5.54)

\[ \pi^\dagger(x) = \int \frac{d^3p}{(2\pi)^3} \left( \frac{E_p}{2} \right)^{1/2} \left[ b_p e^{ix\cdot p} - c_p^\dagger e^{-ix\cdot p} \right] \] (5.55)

with the commutation relations

\[ [\psi(x), \pi(y)] = i\delta^{(3)}(x-y) \]

all other commutators are vanishing

\[ 0 = [\psi(x), \pi(y)^\dagger] = [\psi(x), \psi(y)] = [\pi(x), \pi(y)] = \ldots \]

You can check these commutation relations imply

\[ [b_p, b_q^\dagger] = (2\pi)^3 \delta^{(3)}(p-q) \] (5.56)

\[ [c_p, c_q^\dagger] = (2\pi)^3 \delta^{(3)}(p-q) \] (5.57)

and the rest vanishes

\[ 0 = [b_p, b_q] = [c_p^\dagger, c_q^\dagger] = [b_p, c_q] = [b_p^\dagger, c_q^\dagger] \]

naturally including their hermitian conjugates, e.g. \([c_p, c_q] = 0\).

**Interpretation:** The quantization of a complex scalar field gives rise to two types of creation operators \(b_p^\dagger\) and \(c_q^\dagger\). These are particles and anti-particles.

The complex scalar field has an internal \(U(1)\) symmetry, which as we saw in Sec. 4.2.2 gives rise to a conserved current

\[ j^\mu = i(\partial^\mu \psi^* )\psi - i\psi^*(\partial^\mu \psi) \] (5.58)

and therefore a conserved quantum number

\[ Q = i \int d^3x \left( \psi^* \dot{\psi} - \psi \dot{\psi}^* \right) = i \int d^3x \left( \pi \dot{\psi} - \psi \dot{\psi}^* \right) \] (5.59)

The normal ordered quantum operator is

\[ : Q : \equiv i \int d^3x : (\pi \psi - \psi^* \pi^*) : \]

\[ = - \int d^3x \int \frac{d^3p}{(2\pi)^3} \int \frac{d^3q}{(2\pi)^3} \sqrt{\frac{E_q}{4E_p}} \left[ b_p e^{ix\cdot p} + c_p^\dagger e^{-ix\cdot p} \right] \left[ b_q^\dagger e^{-ix\cdot q} - c_q e^{ix\cdot q} \right] : + h.c. \]

\[ = - \int \frac{d^3p}{(2\pi)^3} \left\{ \left[ b_p^\dagger b_p - c_p c_p^\dagger \right] - \left[ b_p c_p - b_p^\dagger c_p^\dagger \right] + \left[ b_p c_p - b_p^\dagger c_p^\dagger \right] \right\} \]

where we’ve used that hermitian conjugation of the first bracket leaves it unchanged with normal-ordering. The second and third bracket cancel. We finally obtain for the normal-ordered charge operator

\[ : Q : \equiv \int \frac{d^3p}{(2\pi)^3} \left[ c_p^\dagger c_p - b_p^\dagger b_p \right] = N_c - N_b \] (5.60)
which counts the number of anti-particles (created by \( c_p^\dagger \)) minus the number of particles (created by \( b_p^\dagger \)). We know that\(^{23}\)

\[
[H,Q] = 0
\]

which tells us that \( Q \) is a conserved quantity. We saw that in the free field theory this is trivially true with Eq. (5.44), since \([N,H] = 0\). Later, we will see that in interacting field theories, new particles can be created, e.g. \( \mathcal{L} \sim \phi^3 \) can branch a particle into two new ones: \( N_c \) and \( N_b \) will not be individually conserved. However, \( Q = N_c - N_b \) survives as a conserved quantity.

### 5.6 Time dependent operators: the Heisenberg picture

We have been working in the Schroedinger picture, where the operators depend on space but not on time. We will move to the Heisenberg picture and make the operators time-dependent in the usual way

\[
\phi_H(x) = \phi_H(x,t) = e^{iHt} \phi_S(x) e^{-iHt}
\]

and

\[
\pi(x) = e^{iHt} \pi(x) e^{-iHt}
\]

In general

\[
\mathcal{O}_H(t) = e^{iHt} \mathcal{O}_S e^{-iHt}
\]

and so

\[
\frac{d}{dt} \mathcal{O}_H(t) = i[H, \mathcal{O}_H(t)]
\]

In field theory we will drop the subscripts "s" or "H" and indicate using the arguments if we are in the Schrödinger picture \( \phi(x) \) or in the Heisenberg picture \( \phi(x,t) = \phi(x) \).

We assume that at \( t = 0 \) the two pictures coincide and we find that Eq. (5.1) become equal-time commutation relations

\[
[\phi_a(x,t), \pi^b(y,t)] = i\hbar \delta^a_3 \delta^b_3(x-y)
\]

(5.61)

\[
[\phi_a(x,t), \phi_b(y,t)] = 0
\]

(5.62)

\[
[\pi^b(x,t), \pi^a(y,t)] = 0
\]

(5.63)

We can now discuss the time-evolution of the operator \( \phi(x,t) \), e.g.\(^{24}\)

\[
\partial_t \phi(y) = i[H(t), \phi]
\]

\[
= \frac{i}{2} \left[ \int d^3x \pi(x,t)^2 + (\nabla \phi(x,t))^2 + m^2 \phi(x,t)^2, \phi(y) \right]
\]

\[
= i \int d^3x \pi(x)(-i)\delta^{(3)}(x-y)
\]

\[
= \pi(y)
\]

\(^{23}\) Ex: Show this!

\(^{24}\) We use

\[
\]

if the commutator is just a \( c \)-valued function

\[
[A, B] = f
\]

e.g. \( f = \delta(x-y) \).
And
\[ \partial_t \pi(y) = i[H, \pi] = \frac{i}{2} \left[ \int d^3x \pi(x) \nabla_x \phi(x)^2 + m^2 \phi(x)^2, \pi(y) \right] \]
\[ = i \int d^3x \nabla_x [\pi(x), \pi(y)] \nabla_x \phi(x) + im^2 \phi(x) \delta^3(x - y) \]
\[ = - \left( \int d^3x (\nabla_x \delta^3(x - y)) \nabla_x \phi(x) \right) - m^2 \phi(y) \]
\[ = \nabla_y^2 \phi(y) - m^2 \phi(y) \]
which when combined with \( \partial_t \phi(y) = \pi(y) \) gives the Klein-Gordon equation
\[ \partial_{\mu} \partial^{\mu} \phi + m^2 \phi = 0 \]
We find the important result: Field operators satisfy the same equation as the classical field. This starts to resemble a relativistic theory.

How does the Heisenberg field operator expansion look like?
We want to transform \( \phi_S(x) \rightarrow \phi(x) = e^{iHt} \phi_S(x) e^{-iHt} \)
and therefore need to transform the creation and annihilation operators in the expansion
\[ e^{iHt} a_p e^{-iHt} = e^{-iE_p t} a_p \]
\[ e^{iHt} a_p^\dagger e^{-iHt} = e^{+iE_p t} a_p^\dagger \]
which we can prove using the Baker-Campbell-Hausdorff relation
\[ e^X Y e^{-X} = \sum_{m=0}^{\infty} \frac{1}{m!} [X, Y]_m \]
with the general commutator
\[ [[X, Y]]_m = [X, [X, Y]]_{m-1} \quad \text{and} \quad [X, Y]_0 = Y \]
We know that \( [H, a_p] = -E_p a_p \), and so \( [iHt, a_p]_m = (-iE_p t)^m a_p \)
which is all we need to show
\[ e^{iHt} a_p e^{-iHt} = \sum_{m=0}^{\infty} \frac{1}{m!} [iHt, a_p]_m = e^{-iE_p t} a_p \]
An analogous calculation (or hermitian conjugation) allows us to obtain the Heisenberg transformation of \( a_p^\dagger \).
Plugging this into Eq. (5.9) means that the field operator of the real scalar in the Heisenberg picture \( \phi(x, t) = e^{iHt} \phi(x) e^{-iHt} \) has the following expansion
\[ \phi(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \left[ a_p e^{-ip \cdot x} + a_p^\dagger e^{ip \cdot x} \right] \]
where the exponent has the opposite sign, because we have used the Lorentz-vectors
\[ p \cdot x = p_\mu x^\mu = E p t - p \cdot x \]
You can easily check that Eq. (5.67) indeed satisfies the Klein-Gordon equation with \( E_p^2 = p^2 + m^2 \).

For completeness, we show the straight-forward result for the complex scalar field of Eq. (5.52) in the Heisenberg picture

\[
\psi(x) = \int \frac{d^3 p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \left[ b_p e^{-ip \cdot x} + c_p^\dag e^{ip \cdot x} \right] \\
\psi^\dag(x) = \int \frac{d^3 p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \left[ b_p^\dag e^{ip \cdot x} + c_p e^{-ip \cdot x} \right] \tag{5.68} \tag{5.69}
\]

5.7 Causal quantum fields

We already know from Eq. (5.61) that \( \phi(x) \) and \( \pi(x) \) satisfy equal time commutation relations, e.g.
\[ [\phi(x, t), \pi(y, t)] = i \delta^{(3)}(x - y) \tag{5.70} \]

We would like to study now: how does the commutator look like at arbitrary space-time distances? This brings us back to the discussion of causality in the introduction. Recall that we must require that space-like separated events \( (x - y)^2 < 0 \) do not overlap. Or in other words, we should be able to find a basis in which matrix elements of operators can be brought into diagonal form (for \( x \) and \( y \) space-like).

This is equivalent to requiring that space-like separated operators commute
\[
[\mathcal{O}_1(x), \mathcal{O}_2(y)] = 0 \quad \text{if } (x - y)^2 < 0 \tag{5.71}
\]

which ensures that a measurement at \( x \) will not affect a measurement at \( y \) if they are not causally connected. Let us check if we satisfy this property. We define
\[
\Delta(x - y) \equiv [\phi(x), \phi(y)] \tag{5.72}
\]

By direct substitution we find that in the free theory the \( \Delta(x - y) \) is simply a \( c \)-number function with the integral expression\(^\text{25}\)
\[
\Delta(x - y) = \int \frac{d^3 p}{(2\pi)^3} \frac{1}{2E_p} \left[ e^{-(x-y) \cdot p} - e^{i(x-y) \cdot p} \right] \tag{5.73}
\]

Plugging the expression for \( \phi(x) \) from Eq. (5.67) into the definition of \( \Delta(x - y) \) yields
\[
\Delta(x - y) \equiv [\phi(x), \phi(y)] = \int \frac{d^3 p}{(2\pi)^3} \int \frac{d^3 q}{(2\pi)^3} \frac{1}{4E_pE_q} \left[ a_p e^{-ip \cdot x} + a_p^\dag e^{ip \cdot x}, a_q e^{-iq \cdot y} + a_q^\dag e^{iq \cdot y} \right] \tag{5.74}
\]
\[
= \int \frac{d^3 p}{(2\pi)^3} \int \frac{d^3 q}{(2\pi)^3} \frac{1}{4E_pE_q} \left\{ [a_p, a_q] e^{-ip \cdot x -iq \cdot y} + [a_p^\dag, a_q^\dag] e^{ip \cdot x +iq \cdot y} \right. \\
\left. + [a_p^\dag, a_q] e^{ip \cdot x -iq \cdot y} + [a_p, a_q^\dag] e^{ip \cdot x +iq \cdot y} \right\} \tag{5.75}
\]

\(^{25}\) Thanks to Nepomuk Ritz for sending in the solution!
Using the commutation relations

\[ [a_p, a_q] = 0 \quad \quad [a_p, a_q^\dagger] = (2\pi)^3 \delta^{(3)}(p - q) \]

\[ [a_p^\dagger, a_q] = 0 \quad \quad [a_p^\dagger, a_q^\dagger] = -(2\pi)^3 \delta^{(3)}(q - p). \]

this expression simplifies to

\[ \Delta(x - y) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_p} \left[ e^{-i(x-y)\cdot p} - e^{i(x-y)\cdot p} \right] \]

which is the desired result.

**Properties of \( \Delta(x - y) \):**

1. \( \Delta(x - y) \) is Lorentz-invariant. The exponentials are functions of contracted four-vectors and we have just shown the integration measure \( \int d^3p/(2E_p) \) to be Lorentz-invariant.

2. For time-like distances \( \Delta(x - y) \) does not vanish: if we take \( z^\mu = x^\mu - y^\mu = (t, 0, 0, 0)^\mu \), we obtain

\[ \Delta(z) = [\phi(x, t), \phi(x, 0)] \sim \int d^3p \frac{1}{2E_p} \sin(E_p t) \]

which is non-vanishing.

3. For space-like distances \( \Delta(x - y) \) vanishes. We first show it for a space-like distance at equal times

\[ (x - y)^2 = (t - t)^2 - (x - y)^2 = -(x - y)^2 < 0 \]

We obtain in this case

\[ [\phi(x, t), \phi(y, t)] = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2\sqrt{p^2 + m^2}} \left[ e^{i(x-y)\cdot p} - e^{-i(x-y)\cdot p} \right] \]

This vanishes. We can e.g. replace \( p \to -p \) in the second term since it is just an integration variable and we are integrating over a symmetric interval, which changes the sign in the exponent and causes the sum to identically vanish. Since \( \Delta(x - y) \) is Lorentz-invariant, it only depends on \( (x - y)^2 \) and must therefore vanish for all \( (x - y)^2 < 0 \) (and not only for equal times).

We conclude that we have defined a theory of **causal quantum fields**, which solves this problem of relativistic quantum mechanics as described in the introduction. This property will continue to hold in interacting theories.

### 5.7.1 Propagators in quantum field theory

We have already seen the usefulness of the Green’s function method in Sec. 4.3. We will encounter a similar object in quantum field theories. Continuing with the causality problem, we could ask the question: what is the amplitude to create a particle at \( x \) and observe it at \( y \) while leaving the vacuum \( |0⟩ \) undisturbed? We calculate

Even though \( [\phi(x), \phi(y)] \) will not be a \( \mathcal{C} \)-function anymore, but will be operator valued.

Consider a purely space-like distance:

\( x^0 - y^0 = 0 \) and \( x - y = r \) and repeat steps after Eq. (4.92)

\[ D(x - y) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_p} e^{-ipr} (x - y) \]

\[ = \frac{2\pi}{(2\pi)^3} \int_0^\infty dp \frac{p^2}{2\sqrt{p^2 + m^2}} e^{ipr} e^{-ipr} \]

\[ = -\frac{1}{2(2\pi)^2 r} \int_{-\infty}^{\infty} dp \frac{pe^{ipr}}{\sqrt{p^2 + m^2}} \]

The integrand can be considered a complex function with branch-cuts on the imaginary axis (from the \( \sqrt{p^2 + m^2} \)) starting at \( \pm i m \). We evaluate
\[ \langle 0 | \phi(x) \phi(y) | 0 \rangle = \int \frac{d^3p}{(2\pi)^3} \frac{d^3q}{(2\pi)^3} \frac{1}{\sqrt{4E_p E_q}} \langle 0 | a_p a_q^\dagger | 0 \rangle e^{-ip \cdot x + iq \cdot y} \]
\[ = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_p} e^{-ip \cdot (x - y)} \]
\[ \equiv D(x - y) \]  
(5.74)

We call \( D(x - y) \) the propagator. For space-like separations \((x - y)^2 < 0\), we can show, see Eq. (5.73), that
\[ D(x - y) \sim e^{-m|x - y|} \]

Although it decays exponentially quickly outside the light-cone, it is non-vanishing! The quantum field leaks out of the light-cone, violating causality? How do we reconcile this with the fact, Eq. (5.71), that all space-like separated measurements commute? We can write the commutator Eq. (5.72) as

\[ [\phi(x), \phi(y)] = D(x - y) - D(y - x) = 0 \quad \text{if} \quad (x - y)^2 < 0 \]  
(5.75)

Let us describe what this equation means: we know that for space-like separation \((x - y)^2 < 0\) of \(x\) and \(y\), there is no Lorentz invariant ordering of events: if a particle can travel from \(x \rightarrow y\), it can just as easily travel from \(y \rightarrow x\). In any measurement, the amplitudes for the two processes cancel.

Let us consider this for a complex scalar-field, which has distinct particle and anti-particle excitations, see Eq. (5.52). When the complex scalar field \(\psi(x)\) is quantized,

- \(\psi(x)\) will create positive charged particles and destroy negatively charged ones
- \(\psi^\dagger(x)\) will create negative charged particles and destroy positively charged ones

The commutator \([\psi(x), \psi^\dagger(y)]\) will be generally non-zero but most delicately cancel outside of the light-cone to preserve causality.

The interpretation of the two terms in Eq. (5.75) now has charges attached. Since
\[ \psi(x)|0\rangle \sim (\ldots) c_p^\dagger |0\rangle \]
produces an anti-particle at \(x\) and
\[ \psi^\dagger(y)|0\rangle \sim (\ldots) b_p^\dagger |0\rangle \]
produces a particle at \(y\), we can interpret the commutator
\[ [\psi(x), \psi^\dagger(y)] = D(x - y) - D(y - x) \]
\[ = \langle 0 | \psi(x) \psi^\dagger(y) | 0 \rangle - \langle 0 | \psi^\dagger(y) \psi(x) | 0 \rangle \]
\[ = "\text{particle } y \rightarrow x" - "\text{anti-particle } x \rightarrow y" \]

In order for these two processes to cancel, both particles must exist and they must have the same mass.
In quantum field theory, causality requires that every particle has a corresponding anti-particle with the same mass and opposite quantum numbers (here it would be electric charge). What is the anti-particle of the real Klein-Gordon field? The particle of the real field is its own anti-particle.\(^{26}\)

5.8 The Stückelberg-Feynman propagator

We will now discover one of the most important quantities in interacting field theories, the Stückelberg-Feynman propagator. We will define the product of fields so that an interpretation in terms of the causal propagation of particles is always possible

\[
\Delta_F(x - y) = \langle 0 | T\phi(x)\phi(y) | 0 \rangle = \begin{cases} 
D(x - y) & x^0 > y^0 \\
D(y - x) & y^0 > x^0 
\end{cases}
\]

(5.76)

where \(T\) is the time ordering, which places all operators evaluated at later times to the left

\[
T\phi(x)\phi(y) = \begin{cases} 
\phi(x)\phi(y) & x^0 > y^0 \\
\phi(y)\phi(x) & y^0 > x^0 
\end{cases}
\]

(5.77)

Claim: We can express the Feynman propagator of Eq. (5.76) in terms of a Lorentz-invariant expression

\[
\Delta_F(x - y) = \int_{\text{contour}} \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2} e^{-ip \cdot (x - y)}
\]

The meaning of the "contour" will become clear soon. This is the first integral over 4-momenta. Until now, we have integrated only over 3-momenta and kept the \(p^0\) on the mass shell with \(p^0 = E_p\). Here, we have no such condition. The integral for \(\varepsilon = 0\) is ill-defined, because for each \(p\) the denominator has a pole at

\[
p^2 - m^2 = (p^0)^2 - p^2 - m^2 \implies p^0 = \pm E_p = \pm \sqrt{p^2 + m^2}
\]

We need a prescription to avoid these singularities in the \(p^0\) integration. The Feynman propagator follows from the contour

Figure 5.2: The contour \(\mathcal{C}\) for the Feynman propagator in the complex \(p^0\)-plane.

\(^{26}\) And by the property of anti-particles of having opposite quantum numbers, we found an alternative way to show that real scalar fields are neutral.
Proof: We first rewrite the numerator
\[ \frac{1}{p^2 - m^2} = \frac{1}{(p^0)^2 - E_p^2} = \frac{1}{(p^0 - E_p)(p^0 + E_p)} \]
The residue is
\[ \operatorname{Res}_{p^0 = \pm E_p} \left[ \frac{1}{(p^0 - E_p)(p^0 + E_p)} \right] = \pm \frac{1}{2E_p} \]

For \( x^0 > y^0 \): we can close the contour in the lower half plane \( p^0 = -i|\rho| \), since here
\[ e^{-ip^0(x^0 - y^0)} = e^{-|\rho|(x^0 - y^0)} \to 0 \]
for \( p^0 \to -i\infty \) and therefore the radial integration will not contribute. We pick up the residue at \( p^0 = +E_p \) which gives
\[ \oint_C dp^0 \frac{i}{p^2 - m^2} e^{-ip^0(x^0 - y^0)} = -2\pi i \frac{1}{2E_p} e^{-iE_p(x^0 - y^0)} \]
where the minus sign is due to us taking a clockwise contour. Therefore when \( x^0 > y^0 \), we get
\[ \Delta_F(x - y) = \int \frac{d^3p}{(2\pi)^3} \frac{2\pi i}{2E_p} e^{+iE_p(x^0 - y^0) + ip(x - y)} \]
\[ = \frac{2\pi}{(2\pi)^3} \frac{1}{2E_p} e^{-ip(x - y)} \]
\[ = D(x - y) \]
which is the Feynman propagator for \( x^0 > y^0 \).

For \( x^0 > y^0 \): we will close the contour in the upper-half plane. We catch the residue at \(-E_p\) in an anti-clockwise direction and find
\[ \Delta_F(x - y) = \int \frac{d^3p}{(2\pi)^3} \frac{2\pi i}{(-2E_p)} e^{-iE_p(y^0 - x^0) - ip(y - x)} \]
\[ = \frac{2\pi}{(2\pi)^3} \frac{1}{2E_p} e^{-ip(y - x)} \]
\[ = \frac{2\pi}{(2\pi)^3} \frac{1}{2E_p} e^{-ip(y - x)} \]
where in the step from the 2nd to the 3rd line, we have flipped the integration variable \( p \to -p \) which is valid since the rest of the expression (including the boundaries) is symmetric in \( p \). Also in this case we reproduce the Feynman propagator. QED.

Instead of explicitly showing the contour, we usually write for the Feynman propagator the beautiful expression
\[ \Delta_F(x - y) = \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2 + i\varepsilon} e^{-ip(x - y)} \]  
(5.78)
with \( \varepsilon > 0 \) and infinitesimal. This shifts the poles slightly off the real axis, such that the integral along \( p^0 \) is equivalent to the contour \( C \). This form of the propagator is called the \( i\varepsilon \)-prescription.

Recall the residue for a simple pole is
\[ \operatorname{Res}_{z = c} f(z) = \lim_{z \to c} (z - c)f(z) \]
and that for a positively oriented simple closed curve \( C \)
\[ \oint_C f(z)dz = 2\pi i \sum \operatorname{Res}_{z = c_i} f(z) \]
where the sum is over the residues which are enclosed by \( C \). Clockwise encircled residues contribute with the opposite sign.

This is the most important result of this chapter.
\(-E_p + i \epsilon\) \quad \Re(p^0) \\
\bullet \\
\Im(p^0) \\
\bullet \\
\,+E_p - i \epsilon

\[ \begin{align*}
\mathcal{G}(p^0) & = E_p + i \epsilon \\
\mathcal{R}(p^0) & = -E_p + i \epsilon
\end{align*} \]

Figure 5.3: The \(i\epsilon\)-prescription for the Feynman contour. Note that the \(\epsilon\) is not exactly the same as in the propagator, they are proportional to each other since \(\epsilon_{\text{prop}} = 2\epsilon_{\text{plot}}E_p\)

5.8.1 Green’s Functions

We will now see that the objects we’ve been discussing above are closely related to the Green’s functions which we have discussed in Sec. 4.3. In fact, if we stay away from the singularities, we obtain

\[
(\Box + m^2)\Delta_F(x - y) = \int \frac{d^4 p}{(2\pi)^4} \frac{i}{p^2 - m^2} e^{-ip(x-y)} \]

\[
= -i \int \frac{d^4 p}{(2\pi)^4} e^{-ip(x-y)} \]

\[
= -i \delta^{(4)}(x - y) \quad (5.79)
\]

Note, that we did not need the contour in the derivation. For other applications it is useful to pick alternative contours, which also result in Green’s functions.

We can define a **retarded** Green’s function \(\Delta_R(x - y)\) with

\[
\Delta_R(x - y) = \begin{cases} 
D(x - y) - D(y - x) & x^0 > y^0 \\
0 & y^0 > x^0 
\end{cases}
\]

which corresponds to the contour in Fig. 5.4 (left) This is useful if we know the initial value of a field configuration and want to know what it evolves into in the presence of a source, e.g.

\[
\Box \phi + m^2 \phi = J(x)
\]

Along the same lines, we can define an **advanced** Green’s function, see contour in Fig. 5.4 (right)

\[
\Delta_A(x - y) = \begin{cases} 
0 & x^0 > y^0 \\
-D(x - y) + D(y - x) & y^0 > x^0 
\end{cases}
\]

which is helpful if we know the final field configuration and want to figure out where it came from.

Figure 5.4: Contours for the retarded and advanced Green’s function in the complex \(p^0\)-plane.
Using Heaviside $\theta$-functions we can express the three types of Green’s functions as:

$$
\Delta_F(x - y) = \langle 0 | T \phi(x)\phi(y) | 0 \rangle \\
= \theta(x^0 - y^0) \langle 0 | \phi(x)\phi(y) | 0 \rangle + \theta(y^0 - x^0) \langle 0 | \phi(x)\phi(y) | 0 \rangle
$$

and

$$
\Delta_R(x - y) = \theta(x^0 - y^0) \langle 0 | [\phi(x), \phi(y)] | 0 \rangle \\
\Delta_A(x - y) = -\theta(y^0 - x^0) \langle 0 | [\phi(x), \phi(y)] | 0 \rangle
$$

We are still a long way from being able to do any real calculation, since so far we have only talked about the free Klein-Gordon theory, where the field equations are linear and there are no interactions.

On the other hand, the formalism here is extremely important since the free theory forms the basis for doing perturbative calculations in interacting theories.

### 5.8.2 Particle creation by a classical source

Even though we have no interactions yet, we can discuss what happens if we disturb the field by a classical, external source $J(x)$

$$(\Box + m^2)\phi(x) = J(x) \tag{5.80}$$

which follows from the Lagrangian

$$
L = \frac{1}{2} \partial^\mu \phi \partial_\mu \phi - \frac{1}{2} m^2 \phi^2 + J \cdot \phi \tag{5.81}
$$

We will now turn on $J(x)$ for a finite time. Before we turn on $J(x)$, the field has the form of a free field, see Eq. (5.67)

$$
\phi_0(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \left[ a_p e^{-ip \cdot x} + a_p^\dagger e^{ip \cdot x} \right]
$$

without a source, this would be the solution for all time. We can construct the solution with a source, by using the retarded Green’s function

$$(\Box + m^2)\Delta_R(x - y) = -i\delta^{(4)}(x - y) \tag{5.82}$$

$$\Delta_R(x - y) = 0, \quad (x^0 < y^0) \tag{5.83}$$

The second requirement, that $\Delta_R$ be the retarded Green function, is required so that the boundary condition $\phi(x) \to \phi_0(x)$ as $x_0 \to -\infty$ is satisfied.

$$
\phi(x) = \phi_0(x) + i \int d^4 y \Delta_R(x - y) J(y)
= \phi_0(x) + i \int d^4 y \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_p} \theta(x^0 - y^0) \left[ e^{-ip \cdot (x-y)} - e^{ip \cdot (x-y)} \right] J(y)
$$

We have used the fact that if we wait until all of $J(y)$ is in the past, the $\theta(x^0 - y^0)$ equals 1 over the whole domain of integration.

You can check that this is the correct Ansatz, by plugging the first line into Eq. (5.80) and using the properties of the retarded Green’s function

$$
\phi(x) = (\Box + m^2)\phi_0(x) + i \int d^4 y (\Box + m^2) \Delta_R(x - y) J(y)
= 0 + i(-i) J(x) = J(x)
$$

which clearly satisfies Eq. (5.80).
and may be dropped. Then the solution involves only the Fourier transform of $J(y)$

$$J(p) = \int d^4y e^{ip\cdot y} J(y)$$

evaluated at $p^\mu$ such that

$$p^2 = m^2 \quad \text{(on-shell)} \quad (5.84)$$

and with $(J(p))^* = J(-p)$ because $J(y)$ is real, we obtain

$$\phi(x) \overset{x^0 \rightarrow +\infty}{=} \phi_0(x) + i \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_p} \left[ e^{-ip\cdot x} J(p) - e^{ip\cdot x} (J(p))^* \right]$$

We can now group positive frequency terms together with $a_p$ and negative frequency terms together with $a_p^*$:

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \left[ \left( a_p + \frac{i}{\sqrt{2E_p}} J(p) \right) e^{-ip\cdot x} + \left( a_p^* - \frac{i}{\sqrt{2E_p}} (J(p))^* \right) e^{ip\cdot x} \right]$$

Since all observables are built out of the fields, we have solved the theory. We can now guess (or derive) the Hamiltonian after $J(x)$ has been switched on and off. The free Hamiltonian (with $J(x) \equiv 0$) is

$$H = \int \frac{d^3p}{(2\pi)^3} E_p a_p^* a_p$$

and

$$\langle 0 | H | 0 \rangle = 0$$

If we act with $J(y)$ and study the system in the far future ($x^0 \rightarrow +\infty$), we can get $H_J$ looking at the above discussion of the retarded Green’s function by just replacing

$$a_p \rightarrow a_p + \frac{i}{\sqrt{2E_p}} J(p)$$

$$a_p^* \rightarrow a_p^* - \frac{i}{\sqrt{2E_p}} (J(p))^*$$

to obtain

$$H_J = \int \frac{d^3p}{(2\pi)^3} E_p \left( a_p^* - \frac{i}{\sqrt{2E_p}} (J(p))^* \right) \left( a_p + \frac{i}{\sqrt{2E_p}} J(p) \right)$$

The energy of the system in the far future after the source has been switched off, is therefore

$$\langle 0 | H_J | 0 \rangle = \int \frac{d^3p}{(2\pi)^3} E_p \frac{|J(p)|^2}{2E_p} \quad (5.85)$$

Note that because we are in the Heisenberg representation, we are still in the ground state of the free theory – the state has not evolved.

How can we interpret Eq. (5.85)? In the far future, we are in the free theory again and the spectrum of the Hamiltonian is just free.
and each Fourier component of \( J(p) \) produces particles\(^{27} \) with a probability density \( |J(p)|^2 / (2E_p) \) for creating a particle in the momentum eigenstate \( p \). The total number of particles produced is

\[
\int dN = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_p} |J(p)|^2
\]

We see from the condition Eq. (5.84), that only the Fourier modes of \( J(x) \) which are in resonance \( (p^2 = m^2) \) with the on-shell Klein-Gordon waves are effective at creating particles. This is just the classical phenomenon of resonance occurring in the quantum field theory setting.

### 5.9 Non-relativistic fields

We will now derive the non-relativistic limit of the Klein-Gordon equation to see if we can recover the Schrödinger equation. Let us start with a classical complex scalar field which satisfies the Klein-Gordon equation. We decompose

\[
\psi(x, t) = e^{-imt} \tilde{\psi}(x, t)
\]

Plugging this into the Klein-Gordon equation

\[
0 = (\partial_t^2 - \nabla^2 + m^2) \psi = e^{-imt} \left[ (\partial_t^2 - 2im\partial_t - \nabla^2) \tilde{\psi}(x, t) \right]
\]

where the \( m^2 \) term cancelled against the time derivatives. In the non-relativistic limit

\[
|p| \ll m
\]

and so after a Fourier transformation\(^{28} \)

\[
\partial_t \tilde{\psi}(x, t) \rightarrow -iE_p \tilde{\psi}(p), \quad \partial_t \psi(x, t) \rightarrow -i(E_p - m) \psi(p)
\]

in the non-relativistic limit, \( E_p - m \ll m \) and we find

\[
|\partial_t^2 \tilde{\psi}(x, t)| \ll m|\partial_t \tilde{\psi}(x, t)|.
\]

We therefore drop the term with two time derivatives to obtain

\[
\frac{i}{m} \frac{\partial \tilde{\psi}}{\partial t} = -\frac{1}{2m} \nabla^2 \tilde{\psi}
\]

This looks like the Schrödinger equation, except it does not have a probabilistic interpretation – it is just a free classical field, evolving according to an equation of first order in the time derivatives.

In Sec. 4.1.4, we discussed a Lagrangian which was first order in the time derivatives. Once again, we can derive Eq. (4.23) from

\( ^{27} \) This can be made more explicit once we discuss interacting fields and have developed the relevant formalism. Here you are asked to intuit this interpretation by starring long enough at the form of the free Hamiltonian in the absence of a source.

\( ^{28} \) Recall

\[
\phi(x) = \int \frac{d^3p}{(2\pi)^3} \left[ \phi(p) e^{-ip \cdot x} \right]
\]

with \( p^2 = m^2 \), such that \( (\Box + m^2)\phi = 0 \).
the relativistic Klein-Gordon Lagrangian using $\partial_t \psi \ll m\psi$

$$\mathcal{L} = \psi^*(\Box - m^2)\psi$$

$$= |\dot{\psi}|^2 - |\nabla \psi|^2 - m^2|\psi|^2$$

$$= | - im\dot{\psi} + \partial_x \psi|^2 - |\nabla \dot{\psi}|^2 - m^2|\dot{\psi}|^2$$

$$\approx im \left( \dot{\psi}^*(\partial_t \dot{\psi}) - \dot{\psi} (\partial_t \dot{\psi}^*) \right) - |\nabla \dot{\psi}|^2$$

Once we divide by $1/2m$, replace $\tilde{\psi} \rightarrow \psi$ and partially integrate in time, we recover Eq. (4.23) which also was first order in the time derivatives

$$\mathcal{L} = i\psi^*(\partial_t \dot{\psi}) - \frac{1}{2m}(\nabla \psi^*)(\nabla \psi) \quad (5.89)$$

Like its relativistic origin, this Lagrangian has a $U(1)$ symmetry $\psi \rightarrow e^{i\alpha}\psi$. The associated conserved current

$$j^\mu(x) = \begin{pmatrix} \frac{1}{2m}(\psi^* \nabla \psi - \psi \nabla \psi^*) \\ -\psi^* \psi \end{pmatrix} \quad (5.90)$$

Let us compute the Hamiltonian with the goal of quantizing the theory. The conjugate momentum is

$$\pi = \frac{\partial \mathcal{L}}{\partial (\partial_t \psi)} = i\psi^*$$

Interestingly, the momentum does not depend on time derivatives, which is consistent for a theory that only is first order in the time derivatives.\(^\text{29}\) We compute the Hamiltonian

$$\mathcal{H} = \pi \dot{\psi} - \mathcal{L}$$

where the time derivatives drop out. As before we quantize by imposing in the Schrödinger picture the canonical commutation relations

$$[\psi(x), \pi(y)] = i\delta^{(3)}(x-y)$$

or

$$[\psi(x), \psi^\dagger(y)] = \delta^{(3)}(x-y)$$

and all others vanishing

$$[\psi(x), \psi(y)] = [\psi^\dagger(x), \psi^\dagger(y)] = 0$$

We can Fourier expand the field operator $\psi(x)$

$$\psi(x) = \int \frac{d^3p}{(2\pi)^3} a_p e^{ip\cdot x}$$

The commutation relations imply

$$[a_p, a_q^\dagger] = (2\pi)^3 \delta^{(3)}(p-q)$$
As before the vacuum (or ground-state) satisfies \( a_p |0\rangle = 0 \), the multi-particle excitations are proportional to \( a_{p_1} \cdots a_{p_n} |0\rangle \). They are eigenstates of the Hamiltonian

\[
H(p) = \frac{p^2}{2m} |p\rangle
\]

The single particle states satisfy the non-relativistic dispersion relation.

**Comments**

- Quantizing the first order Lagrangian above gives rise to non-relativistic particles of mass \( m \).
- We started with a complex field but found only a single type of particle. The anti-particle is not part of the spectrum. Only with relativity does the existence of anti-particles follow.
- The conserved charge implied by Eq. (5.90) is equal to particle number

\[
Q = \int d^3 x : \psi \dagger \psi :
\]

Even if we introduce interactions \( \Delta L = V(\psi^* \psi) \), particle number will still be conserved. Only with relativity, anti-particles appear and particle number can change.
- We cannot find a non-relativistic limit of a real scalar field. In the relativistic theory the particles were their own anti-particles in this case. There can be no way to construct a multi-particle theory which conserves particle number!

### 5.9.1 A special case: quantum mechanics

How do we recover quantum mechanics? In quantum mechanics, we describe physics in terms of the momentum operator \( P \) and the position operator \( X \). We already have the momentum operator

\[
P = \int \frac{d^3 p}{(2\pi)^3} p a_p \dagger a_p
\]

In the non-relativistic limit we can also construct a position operator, see discussion in Sec. 5.4.1. The operator

\[
\psi \dagger (x) = \int \frac{d^3 p}{(2\pi)^3} a_p \dagger e^{-ip \cdot x}
\]

creates a particle at \( x \) and we write

\[
|x\rangle = \psi \dagger (x) |0\rangle
\]

The best candidate for a position operator is then

\[
X = \int d^3 x \ x \ \psi \dagger (x) \psi (x)
\]
so that

\[ X |x\rangle = x |x\rangle \]

Let us now construct a state \( |\varphi\rangle \) which we would usually call the **Schrödinger wavefunction** (in the position representation). We obtain it by taking superpositions of one-particle states \( |x\rangle \)

\[ |\varphi\rangle = \int d^3x \varphi(x) |x\rangle \]

Let us check if it has the right properties. The position operator has the right action, as we can see

\[ X^i |\varphi\rangle = \int d^3x x^i \varphi(x) |x\rangle \] \hspace{1cm} (5.92)

The momentum operator acts as

\[ P^i |\varphi\rangle = \int d^3x \left( -i \frac{\partial \varphi(x)}{\partial x^i} \right) |x\rangle \] \hspace{1cm} (5.93)

So we learn that when acting on one-particle states, \( P \) and \( X \) act as position and momentum operators in quantum mechanics, with

\[ [X^i, P^j] |\varphi\rangle = i\delta^{ij} |\varphi\rangle \]

What about time-evolution? The wave-function \( \varphi(x,t) \) evolves in time according to the Hamiltonian

\[ H = \int d^3x \frac{1}{2m} \nabla\psi^* \nabla\psi = \int d^3p \frac{p^2}{2m} a_p^\dagger a_p \]

and we find

\[ i \frac{\partial \varphi}{\partial t} = H \varphi = -\frac{1}{2m} \nabla^2 \varphi \] \hspace{1cm} (5.94)

This is the same equation obeyed by the original (full quantum) field. Only this time, it is really the Schrödinger equation with the usual probabilistic interpretation for the wavefunction \( \varphi \). In particular, the **total probability** which is conserved in quantum mechanics arises as the **conserved charge** of the Noether current:

\[ Q = \int d^3x |\varphi(x)|^2 \]

It is useful to know that if we treat the one-particle Schrödinger equation as the equation for a quantum field then it will give the proper generalization to multi-particle quantum field theory. Historically, the fact that the equation for the classical field Eq. (5.88) and the one-particle wavefunction Eq. (5.94) gave rise to considerable confusion and people thought that perhaps we are quantizing the wavefunction itself and called it "second quantization". This is clearly incorrect, we are only quantizing the classical field once!
5.9.2 Non-relativistic interactions

Often we are interested in some fixed background potential $V(x)$. This can be incorporated into field theory by working with a Lagrangian with

$$\mathcal{L} = i \psi^* \left( \partial_t \psi \right) - \frac{1}{2m} (\nabla \psi^*)(\nabla \psi) - V(x)\psi^* \psi$$  \hspace{1cm} (5.95)

We have broken translational symmetry and there will not be an associated, conserved energy-momentum tensor. Such Lagrangians are useful in condensed-matter physics but we almost never encounter them in high-energy physics, where everything is Poincaré invariant, that is translational and Lorentz invariant.

We can also have interactions between particles, which require $n$ particle states with $n \geq 2$. They can arise from Lagrangians of the form

$$\Delta \mathcal{L} = \lambda \psi^*(x)\psi^*(x)\psi(x)\psi(x)$$

which destroys two particles before creating two new ones. In the following chapter we will explore interactions like these in detail for relativistic theories.
6 Interacting quantum fields

We were able to completely solve the free scalar quantum field theory, but nothing interesting happens in this theory. Most importantly, the particle excitations do not interact with each other. Now we will discuss more complicated theories with interaction terms, which in particular will lead to non-linear field equations.

6.1 Dimensional analysis of interactions

We can add the following terms to the free Lagrangian

\[ L = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi^2 - \lambda_3 \phi^3 - \lambda_4 \phi^4 - \lambda_5 \phi^5 + \ldots \]

We call the coefficients of the higher order terms \( \lambda_n \) coupling constants. The equations of motion now contain non-linear terms and they cannot be solved by Fourier analysis as the free Klein-Gordon equation.

When are these additional terms small perturbations? Which terms should we include? What about \( \lambda_15 \phi^{15} \)?

Recall our discussion in Sec. 5.3.1, where we derived the mass dimension of \([\phi] = 1\) and of the action \([S] = 0\) from which we immediately conclude \([L] = 4\). We obtain for the coupling constants

\[ [\lambda_n] = 4 - n, \quad \text{e.g.} \quad [\lambda_3] = 1, \quad [\lambda_4] = 0, \quad [\lambda_5] = -1 \]

So for all but \( \lambda_4 \) we cannot just ask for \( \lambda_n \ll 1 \), since this only makes sense for dimensionless quantities. We will introduce mass scales \( M_i \) for the dimensionful couplings to write

\[ L = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi - \tilde{\lambda}_3 M_3 \phi^3 - \tilde{\lambda}_4 M_4 \phi^4 - \frac{\tilde{\lambda}_5}{M_5} \phi^5 + \ldots \]

Now the statement of \( \tilde{\lambda}_n \ll 1 \) makes sense but it has been accompanied by discussion of the typical energy \( E \) of the process under consideration vs. the mass scale \( M_i \) of the interaction.

We distinguish three classes of interactions

• **Relevant perturbation**: \( \Delta L = -\lambda_3 \phi^3 = -\tilde{\lambda}_3 M_3 \phi^3 \) and \( [\lambda_3] = 1 \). The dimensionless parameter is

\[ \frac{\lambda_3}{E} = \frac{\tilde{\lambda}_3}{E} \frac{M_3}{E} \]

I have restricted the expansion to powers of \( \phi \) but we could have equally added terms of the form

\[ \frac{1}{M^{4-4m-4n}} (\partial_\mu \phi \partial^\mu \phi)^n \phi^m \]

with integer \( n, m \). In particular, we could have added

\[ \Delta L = \frac{1}{M^4} (\partial_\mu \phi \partial^\mu \phi)^2 \]

Here you can see explicitly how the ‘typical energy scale’ \( E \) appears if you replace \( \partial_\mu \sim p_\mu \sim E \). We obtain for the typical size of the interaction \( \sim \frac{E^4}{M^4} \)

that it is clearly an irrelevant perturbation.

Note here, that the meaning of relevant, marginal and irrelevant is tied to the number of space-time dimensions. In \( D = 2 \) for example, \([\phi]=0\) and all the \( \phi^n \) terms are relevant!
What is the meaning of the scale \( E \)? We typically take \( E \) to be the energy scale of the process of interest.\(^1\) This means that when we test the theory at very high energies \( E \gg \lambda_3 \) or \( E \gg M_3 \),\(^2\) \( \lambda_3 \phi^3 \) is a small perturbation, suppressed by \( M_3/E \). If we test at low energies \( E \ll \lambda_3 \) it is a large perturbation. We call terms like this relevant because they are most relevant at low energies, which is where most of the physics we see takes place.

In relativistic theories, \( E > m \), and we can make this a small perturbation by taking \( \lambda_3 \ll m \) or \( M_3 \ll m \).

- **Marginal perturbation:** \( \Delta L = \lambda_4 \phi^4 \) with \([\lambda_4] = 0\). The effects of terms with \( d = 4 \) are independent of energy, and these interactions are called marginal. This term is small if \( \lambda_4 \ll 1 \).

- **Irrelevant perturbation:** \( \Delta L = \lambda_5 \phi^5 \) with \([\lambda_5] = -1 \) (and \( \lambda_n \phi^n \) for \( n \geq 5 \)). The dimensionless parameter is

\[
\lambda_5 = \frac{\tilde{\lambda}_5 E}{M_5}
\]

This interaction is small at low energies \( E \ll M_5 \) and large at high energies. Such perturbations are called irrelevant.

Note, that his naive categorization of relevant, marginal and irrelevant is sometimes subject to change due to quantum corrections.\(^3\)

Further, we see that irrelevant terms can be problematic in QFT since, as we have seen, we often have to integrate to very high energies. These terms lead to so-called ‘non-renormalizable’ field theories and which require some additional work to make sense of the infinities at arbitrarily high energies.

### 6.1.1 Some examples of weakly coupled theories

Here we will only study theories where the interaction terms can be considered small perturbations.

- **\( \phi^4 \) theory:** This is a subset of the marginal and relevant interactions above

\[
\mathcal{L} = \frac{1}{2} \partial_{\mu} \phi \partial^{\mu} \phi - \frac{1}{2} m^2 \phi^2 - \frac{\lambda}{4!} \phi^4
\]

with small \( \lambda \ll 1 \). Is there a reason why we should not have a \( \Delta \mathcal{L} \sim \phi^3 \) coupling? This could be explained by a discrete symmetry \( \phi(x) \rightarrow -\phi(x) \) which allows only even powers in the field.

We can guess the effects of the extra term. If we expand out \( \phi^4 \) in creation and annihilation operators, we find terms like

\[
a_{p_1}^\dagger a_{p_2}^\dagger a_{p_3}^\dagger a_{p_4}, \quad a_{p_1}^\dagger a_{p_2} a_{p_3}^\dagger a_{p_4}, \quad \cdots
\]

These terms create and destroy particles and particle number will in general not be conserved and we could calculate \([H,N] \neq 0\) (even though the discrete parity would be conserved.)

---

\(^1\) Think of the center of mass energy of a scattering, or the mass of particle if you consider decays of particles.

\(^2\) Always assuming that \( \tilde{\lambda}_n = \mathcal{O}(1) \).

\(^3\) One example are the pions of QCD, they are scalars \([\phi] = 1\) but emerge as strongly coupled bound states of two quarks \([\bar{\psi}\psi] = 3\) at low energy. Fermions in a Dirac spinor have mass-dimension \([\psi] = 3/2\).

You might discuss strongly interacting theories in a QFT2 course.
**Why QFT is so simple**

Typically we consider only relevant and marginal couplings, since irrelevant couplings become small at low energies. This simplifies QFT a lot, since out of the infinite number of interaction terms we can write down, only a small number is needed. In the case of a scalar field, we would have

\[
L = \frac{1}{2} \partial_{\mu} \phi \partial^{\mu} \phi - \frac{1}{2} m^2 \phi^2 - \lambda_3 \phi^3 - \lambda_4 \phi^4 + \mathcal{O}(\Lambda/E)
\]

Let us assume, we someday discover the true Lagrangian of the universe which is valid up to very high energies, say \( \Lambda \) with \( [\Lambda] = 1 \). This could be the GUT scale (\( \Lambda \approx 10^{16} \) GeV) or the Planck scale (\( \Lambda \approx 10^{18} \) GeV). If we now investigate nature at the small energies \( E \) accessible to us \( E \ll \Lambda \) and if we assume she is described by a scalar field \( \phi \) (think of the Higgs). This scalar field will then have some complicated couplings \( \lambda_n \phi^n \) which depend on the UV completion of our amazing theory of everything at \( \Lambda \). In terms of the dimensionless couplings \( \tilde{\lambda}_n \), we expect them to scale like

\[
\lambda_n = \frac{\tilde{\lambda}_n}{\Lambda^{n-4}}
\]

Now for experiments at \( E \ll \Lambda \) (think of \( E \approx 10^4 \) GeV at the LHC), the interaction terms of \( \phi^n \) with \( n > 4 \) will be suppressed by very small powers of

\[
\left( \frac{E}{\Lambda} \right)^{n-4}
\]

This simple argument using dimensional analysis tells us that we can usually focus on the first few terms in the interactions. It also tells us that it will be very difficult to figure out the high energy theory. These considerations are a simplified version of what is called effective field theory and Wilson’s renormalization group. Due to the unknown UV physics completing our QFT, we are generally required to write down all the allowed marginal and relevant interactions to capture the IR behaviour of a theory with a scalar field. This discussion is the modern viewpoint of what people previously discussed under the notion of “renormalizable” quantum field theories.

- **Scalar Yukawa theory:** We introduce a real scalar \( \phi \) and a complex scalar \( \psi \) with the Lagrangian

\[
L = \partial_{\mu} \psi^* \partial^{\mu} \psi - M^2 \psi^* \psi + \frac{1}{2} \partial_{\mu} \phi \partial^{\mu} \phi - \frac{1}{2} m^2 \phi^2 - g \psi^* \psi^* \phi
\]

The coupling \( g \) has \( [g] = 1 \) and is a relevant coupling. For the perturbation to be small, we require \( g \ll M, m \). Particle number is again not conserved anymore but since the Lagrangian is invariant under a continuous phase symmetry \((U(1))\)

\[
\psi \rightarrow e^{i\alpha} \psi
\]

a conserved charge \( Q \) exists, with \( [Q, H] = 0 \). The number of \( \psi \)-particles minus the number of \( \bar{\psi} \)-anti-particles is conserved. We will usually denote the antiparticle as \( \bar{\psi} \).

- **Quantum electro-dynamics** This theory involves Dirac 4-spinors\(^4 \) \( \Psi \) and the electro-magnetic vector potential \( A_{\mu} \).

\[
L = \bar{\Psi} (i\gamma^\mu \partial_\mu - m) \Psi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - e \bar{\Psi} \gamma^\mu \Psi A_\mu
\]

You can easily determine the canonical dimensions of the fields to find that \([e] = 0\). You know \( e = -|e| \) as the electron charge. In our units \( |e| \approx 0.3 \) and the interaction term is small.

\(^4\) Which we will discuss in the next chapter.
• **Fermion Yukawa theory** This theory involves spinors $\Psi$ and a real scalar $\phi$

$$
\mathcal{L} = \Psi(i\gamma^\mu \partial_\mu - m)\Psi + \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} M^2 \phi^2 - g \bar{\Psi} \Psi \phi
$$

You can also determine the canonical dimensions of the fields to find that $[g] = 0$. Yukawa invented this theory to describe nucleons ($\Psi$) and pions ($\phi$). Nowadays we use it in the Standard Model (SM), since this is the interaction between the Higgs and the matter fermions (quarks and leptons). Most of the free parameters in the SM are Yukawa coupling constants.

In our now interacting quantum field theory, we impose the equal-time commutation relations

$$
[\phi_a(x, t), \pi^b(y, t)] = i \delta_a^b \delta^{(3)}(x - y)
$$

which are **unaffected by** $\mathcal{L}_{\text{int}}$.\(^5\) You can straight-forwardly check then that the field operator in the Heisenberg picture satisfies the equation of motion including interaction term.

### 6.2 The interaction picture

We have discussed the Schroedinger (states time-dependent, operators not) and the Heisenberg picture (operators time-dependent, states not) in Sec. 5.6 and we will now introduce a hybrid called the **interaction picture**. We write the Hamiltonian as the free theory part $H_0$ plus the interaction $H_{\text{int}}$

$$
H = H_0 + H_{\text{int}}
$$

for the $\phi^4$-example that would be

$$
H = H_{\text{Klein-Gordon}} + \int d^3x \frac{\lambda}{4!} \phi^4(x)
$$

We now define states and operators in the interaction picture as

\[
|\psi(t)\rangle_I = e^{iH_0 t} |\psi(t)\rangle_S \quad \text{(6.3)}
\]

\[
O_I(t) = e^{iH_0 t} O_S e^{-iH_0 t} \quad \text{(6.4)}
\]

when the interaction is small, e.g. $\lambda \ll 1$, then the most important time dependence generated by $H_0$ is already taken care off. The interaction Hamiltonian in the interaction picture is

$$
H_I = (H_{\text{int}})_I = e^{iH_0 t} (H_{\text{int}}) e^{-iH_0 t} \quad \text{(6.5)}
$$

States in the interaction picture evolve according to (starting with the Schroedinger picture result)

$$
\frac{d}{dt} |\psi(t)\rangle_S = H_S |\psi(t)\rangle_S
$$

\(^5\) Note, that if $\mathcal{L}_{\text{int}}$ contained a $\partial_\mu \phi$, the definition of the conjugate momentum $\pi(x)$ would change.
plugging in the definition of the interaction state $|\psi(t)\rangle_I$, we get

$$i\frac{d}{dt}(e^{-iH_0 t} |\psi\rangle) = (H_0 + H_{\text{int}})_S e^{-iH_0 t} |\psi\rangle_I$$

$$i\frac{d}{dt}|\psi\rangle_I = H_I |\psi\rangle_I$$

where we have first canceled $H_0$ on both sides, multiplied the equation by $e^{iH_0 t}$ and used Eq. (6.5).

### 6.2.1 Dyson’s formula

At a fixed time $t_0$, we can expand $\phi(x)$ as before in terms of ladder operators

$$\phi(x, t_0) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \left[ a_p e^{ipx} + a_p^\dagger e^{-ipx} \right]$$

We now want a perturbative expansion that captures the full time-dependence of the quantum field $\phi$. In the Heisenberg picture, we would get

$$\phi_H(x, t) = e^{iH(t-t_0)} \phi_H(x, t_0) e^{-iH(t-t_0)}$$

For $\lambda = 0$, we have $H = H_0$ and this reduces with Eq. (6.4) to

$$\phi(x, t)|_{\lambda=0} = e^{iH_0(t-t_0)} \phi(x, t_0) e^{-iH_0(t-t_0)} = \phi_I(x, t)$$

When $\lambda$ is small, this will give the most important part of the time-dependence. Since $H_0$ is diagonalized by the Fourier expansion, we get

$$\phi_I(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \left[ a_p e^{-ipx} + a_p^\dagger e^{ipx} \right] \bigg|_{x^0 = t-t_0}$$

which is just the familiar result of Eq. (5.67). We now want to express the Heisenberg picture field operator $\phi$ in terms of $\phi_I(x)$, again as before to have the full time-dependence in the operators (while keeping the states time-independent). We can write this as

$$\phi(x, t) = e^{iH(t-t_0)} e^{-iH_0(t-t_0)} \phi_I(x, t) e^{iH_0(t-t_0)} e^{-iH(t-t_0)}$$

$$\equiv U(t, t_0) \phi_I(x, t) U(t, t_0)$$

with the unitary operator

$$U(t, t_0) = e^{iH_0(t-t_0)} e^{-iH(t-t_0)}$$

which is also called the interaction picture propagator or time-evolution operator. States in the interaction picture evolve as

$$|\psi(t)\rangle_I = U(t, t_0)|\psi(t_0)\rangle_I$$

We want to express $U(t, t_0)$ entirely in terms of $\phi_I$ for which we have an explicit expression in terms of ladder operators. We note, that

From the definition of a state in the interaction picture

$$|\psi(t)\rangle_I = e^{iH_0 t}|\psi(t)\rangle_S$$

$$= e^{iH_0 t} e^{-iH_I t}|\psi(0)\rangle_S = U(t, 0)|\psi(0)\rangle_S$$

where we have used the time-evolution in the Schrodinger picture.
$U(t, t_0)$ is the unique solution, with initial condition $U(t_0, t_0) = 1$ of a simple differential equation

$$
\frac{\partial}{\partial t} U(t, t_0) = e^{iH_0(t-t_0)} (H - H_0)e^{-iH(t-t_0)}
$$

$$
= e^{iH_0(t-t_0)} (H_{\text{int}}) e^{-iH(t-t_0)}
$$

$$
= e^{iH_0(t-t_0)} (H_{\text{int}}) e^{-iH_0(t-t_0)} e^{iH_0(t-t_0)} e^{-iH(t-t_0)}
$$

$$
= H_I(t) U(t, t_0)
$$

with $H_I(t)$ as defined in Eq. (6.5). The solution of this equation should naively look something like $U \sim \exp(-i \int^t dt' H_I(t'))$, but that ignores the fact that $[H_I(t), H_I(t')] \neq 0$ when $t \neq t'$.

A more careful treatment shows that the actual solution is

$$
U(t, t_0) = 1 + (-i) \int_{t_0}^t dt_1 H_I(t_1) + (-i)^2 \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 H_I(t_1) H_I(t_2) + \ldots
$$

To verify, that this indeed the correct solution, take the derivative: you find that each term gives the previous one $\times (-i) H_I(t)$. The initial condition is $U(t_0, t_0) = 1$ is trivially satisfied.

We observe that that the factors of $H_I(t_1)$ stand in time order, with the later operators to the left. We can now simplify using the time ordering $T$. Starting with $H_I^2$

$$
\int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 H_I(t_1) H_I(t_2) = \frac{1}{2} \int_{t_0}^t dt_1 \int_{t_0}^t dt_2 T \{ H_I(t_1) H_I(t_2) \}
$$

The time-ordered double integral on the RHS just counts everything twice, since in the $t_1 - t_2$ plane, the integrand $T \{ H_I(t_1) H_I(t_2) \}$ is symmetric about the $t_1 = t_2$ line. Similarly, one can show

$$
\int_{t_0}^t dt_1 \ldots \int_{t_0}^t dt_n H_I(t_1) \ldots H_I(t_n) = \frac{1}{n!} \int_{t_0}^t dt_1 \ldots \int_{t_0}^t dt_n T \{ H_I(t_1) \ldots H_I(t_n) \}
$$

It’s not easy to draw this but you should be able to convince yourself that this is in fact correct. We can now write the time evolution operator in a very compact form:

$$
U(t, t_0) = 1 + (-i) \int_{t_0}^t dt_1 H_I(t_1) + \frac{(-i)^2}{2!} \int_{t_0}^t dt_1 \int_{t_0}^t dt_2 T \{ H_I(t_1) H_I(t_2) \}
$$

$$
+ \frac{(-i)^3}{3!} \int_{t_0}^t dt_1 \int_{t_0}^t dt_2 \int_{t_0}^t dt_3 T \{ H_I(t_1) H_I(t_2) H_I(t_3) \} + \ldots
$$

and we obtain **Dyson’s formula**

$$
U(t, t_0) = T \exp \left( -i \int_{t_0}^t dt' H_I(t') \right)
$$

(6.15)

It is usually very hard (impossible) to compute time ordered exponentials in practice. The power of Dyson’s formula is that just keep the first couple of terms in the expansion when the interaction is a small perturbation.
6.3 A first look at scattering

We will now apply what we have just learned to a the interaction in the scalar Yukawa theory

\[ H_I = g \int d^3x \psi \phi \]

where all the fields are in the interaction picture.\(^7\) This \( H_I \) allows particles to morph into each other. To understand why that is, follow the time evolution of a state

\[ |\psi(t)\rangle_I = U(t, t_0) |\psi(t_0)\rangle_I \]

where \( U(t, t_0) \) is given by Dyson’s formula, which contains an expansion in powers of \( H_I \). Now, \( H_I \) contains annihilation and creation operators for different types of particles. This is a simplified description of meson-nucleon interactions.\(^8\) We can have

- \( \psi^\dagger \sim b^\dagger + c^\dagger \): This operator creates \( \psi \) particles through \( b^\dagger \) and destroys anti-particles through \( c \). Let us call these \( \psi \)-particles nucleons.

- \( \psi \sim b + c^\dagger \): This operator destroys \( \psi \) particles (nucleons) through \( b \) and creates anti-particles (anti-nucleons) through \( c^\dagger \).

- \( \phi \sim a + a^\dagger \): We will call the \( \phi \) particles mesons and \( \phi \) can therefore create and destroy mesons.

Due to the \( U(1) \) symmetry of \( \psi \) in \( \mathcal{L} \), we know that the charge

\[ Q = N_c - N_b \]

will be conserved even in the presence of \( \mathcal{L}_I = -H_I \).

\(^7\) Which means the field operators are in the Heisenberg picture of the free theory.

\(^8\) Nucleons are neutron and the proton, which are composed of down and up quarks. We will now take the historical perspective and consider them elementary particles. In our toy model we will describe them as scalars. In reality nucleons are spin 1/2 particles, and therefore fermions.
The interaction $\psi^\dagger \psi \phi$ contains a collection of creation and annihilation operators, such as

$$c^\dagger b^\dagger a$$

which annihilates a $\phi$ particle (meson) and creates $\psi$ particle and antiparticle (nucleon and anti-nucleon): this corresponds to the decay of the process $\phi \to \bar{\psi} \psi$.

Another example

$$ac^\dagger a \to \psi + \phi \to \psi$$

which corresponds to the absorption of a meson by a nucleon.

At second order in perturbation theory, $\sim H_2^2$, which produces more complicated processes like

$$\bar{\psi} + \psi \to \phi \to \bar{\psi} + \psi$$

nucleon-antinucleon scattering through the creation of an intermediate meson.

### 6.3.1 Scattering simplified

We will now discuss scattering processes. They are particularly convenient since in many cases the initial and final states look like non-interacting particles. What does that mean? In a scattering process, we start with an initial state $|i\rangle$ of isolated particles. The particles are widely separated and do not feel the effects of the interaction: they look like free plane wave states. Once the particles approach each other, they begin to feel the interaction (and the resulting potential) and they evolve in complicated non-linear way as in Eq. (6.14). In this intermediate stage the system will look extremely complicated if expressed in terms of free particles. Particles will be created and destroyed, since in general

$$[H_I, N] \neq 0$$

We will not just have two colliding protons but a very complicated mess of protons, pions, photons, gluons, etc.

The result of the scattering process can have several outcomes. Multiple initial particles could form a bound state. In this case, no matter how long we wait after the scattering, the final state will never look like the eigenstate of the free Hamiltonian $H_0$, because the interaction is responsible for the bound state. The formalism here will not be useful in this case.

We will now consider a process where no bound state is formed. After some long time the system will just be a couple of widely separated (non-interacting) particles. It will look simple again.

Before we go any further, you need to know that this is a bit of a fake. Despite this, our quick and dirty scattering model will still work in the processes we consider here. We can see this by imagining a theory like

$$\mathcal{L} = \mathcal{L}_\phi + \mathcal{L}_\psi - g f(t) \psi^\dagger \psi \phi \quad (6.16)$$

9 Eigenstates of $H_0$ and so eigenstates also of $N$, even though in general $[H, N] \neq 0$

10 e.g. $p + p \to 2D$, two protons forming a deuterium nucleus.

11 Without the interaction the bound state will fly apart.

12 No matter how far in the future we go, we never end up with a collection of free particles. You already know this from electromagnetism: the electron always carries its electromagnetic field with it (which corresponds to a virtual cloud of photons around the electron). Similarly, nucleons always have a cloud of mesons around them. Turning off our interaction, the states will change, and our simple picture is not quite right.
with \( f(t) \) like in Fig. 6.2. The function \( f(t) = 0 \) for large \( |t| \) and \( f(t) = 1 \) around 0 where the scattering happens. A long time after the scattering has occurred, we turn the interaction off very slowly (adiabatically) over a period \( \Delta \); we expect that the simple states in the real theory slowly turn into eigenstates of \( H_0 \) with unit probability.

There will be a one-to-one correspondence between asymptotic simple eigenstates of the full Hamiltonian \( H \) and the eigenstates of the free \( H_0 \). We should recover the full theory if

\[
T \to \infty, \quad \Delta \to \infty, \quad \Delta/T \to 0
\]

where the last limit is required to ensure no edge effects occur (switching on and off of \( H_\text{I} \) is adiabatic). We can justify this approach in a more rigorous manner but we do not have time for the technical details in this course. Look for the LSZ reduction formula in your favorite QFT textbook, named after the German physicists Harry Lehmann, Kurt Symanzik, and Wolfhart Zimmermann.

After this long motivation, we finally get to the meat. We want to solve Eq. (6.7)

\[
\frac{d}{dt} |\psi\rangle = H_\text{I} |\psi\rangle
\]

with the boundary condition

\[
|\psi(-\infty)\rangle = |i\rangle \quad (6.17)
\]

and

\[
|\psi(+\infty)\rangle = |f\rangle \quad (6.18)
\]

The amplitude to go from \(|i\rangle\) to \(|f\rangle\) is

\[
\lim_{t_+ \to +\infty} \langle f|U(t_+, t_-)|i\rangle \equiv \langle f|S|i\rangle \quad (6.19)
\]

where the unitary operator \( S \) is called the S-matrix. Alternatively, we can with Eq. (6.15) write

\[
S = T \exp \left( -i \int d^4x H_\text{I}(x) \right) \quad (6.20)
\]
6.3.2 Example: Meson Decay $\phi \rightarrow \bar{\psi} \psi$

We take the (relativistically) normalized initial $\phi$ and final states $\bar{\psi} \psi$

$$|i\rangle = |p_\phi\rangle = \sqrt{2E_p} a_p^\dagger |0\rangle \quad (6.21)$$

$$|f\rangle = |q_1, q_2\rangle = \sqrt{4E_{q_1} E_{q_2}} b_{q_1}^\dagger c_{q_2}^\dagger |0\rangle \quad (6.22)$$

where $q_{1,2}$ are the momenta of the decay products, a nucleon antinucleon pair ($\bar{\psi} \psi$). We can compute the amplitude for the decay with Eq. (6.15) to leading order in $g$

$$\langle f|S|i\rangle = \langle f|i\rangle - i \langle f| \int_{-\infty}^{\infty} dt H_I(t)|i\rangle + \ldots \quad (6.23)$$

$$= 0 - ig \langle f| \int d^4x \bar{\psi}^\dagger(x) \psi(x) \phi(x)|i\rangle \quad (6.24)$$

We first expand $\phi \sim a + a^\dagger$. We will only need the $a$ piece which annihilates the initial meson state $|i\rangle$ into something $\sim |0\rangle$. We obtain,

$$\langle f|S|i\rangle = i \langle f| \int_{-\infty}^{\infty} dt H_I(t)|i\rangle + \ldots$$

$$= -ig \langle f| \int d^4x \bar{\psi}^\dagger(x) \psi(x) \int \frac{d^4k}{(2\pi)^4} \frac{\sqrt{2E_p}}{\sqrt{2E_k}} a_k a_p^\dagger e^{-ik \cdot x} |0\rangle$$

$$= -ig \langle f| \int d^4x \bar{\psi}^\dagger(x) \psi(x) e^{-ip \cdot x} |0\rangle$$

where we have used the commutation relation of the $a_k$. To get non-zero overlap with $|f\rangle$, only the $b^\dagger$ and $c^\dagger$ contribute, since we need to create a nucleon and anti-nucleon final state from $|0\rangle$. We have,

$$\langle f|S|i\rangle = -ig \langle 0| \int \frac{d^4x d^4k_1 d^4k_2}{(2\pi)^8} \frac{\sqrt{4E_{q_1} E_{q_2}}}{\sqrt{4E_{k_1} E_{k_2}}} e^{-i(p-k_2-k_1) \cdot x} b_{q_1}^\dagger c_{q_2}^\dagger b_{k_1}^\dagger c_{k_2}^\dagger |0\rangle$$

$$= -ig (2\pi)^4 \delta^{(4)}(p - q_1 - q_2) \quad (6.25)$$

and we have found our first quantum field theory amplitude! Notice the $\delta$-function which encapsulates the 4-vector conservation. In particular, the decay will only take place if $m_\phi \geq 2M_\psi$.\footnote{Ex: Show this!}

We post-pone the steps to turn this into a something observable, like a life-time of the meson. You can already see that it might involve some additional insights since once we calculated probabilities, we will get the square of a $\delta$-function.

6.4 Wick’s Theorem

We now want to find a systematic way to compute quantities like

$$\langle f|T\{H_I(x_1) \ldots H_I(x_n)\}|i\rangle$$

from Dyson’s formula. The time-ordering $T$ fixes the order of the operators, but we would have a much simpler calculation if we could move all annihilation operators to the right, where they can remove particles in $|i\rangle$. What we want find is therefore: how to go from time
ordered products to normal order products of fields. This is called **Wick’s theorem.**

We define the **contraction** of two fields

\[
\overline{A(x)B(y)} \equiv T\{A(x)B(y)\} - :A(x)B(y): \quad (6.26)
\]

We can easily see that \(\overline{A(x)B(y)}\) is a \(c\)-number. Let us consider the example of a real scalar field

\[
\phi(x) = \phi^+(x) + \phi^-(x)
\]

decomposed as

\[
\phi^+(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} a_p e^{-ip \cdot x}
\]
\[
\phi^-(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} a_p^\dagger e^{ip \cdot x}
\]

where I would have used the opposite assignment of \(+\) and \(-\) but we have Pauli and Heisenberg to blame.\(^{17}\) Let us consider \(x^0 > y^0\), we obtain

\[
T\{\phi(x)\phi(y)\} = \phi(x)\phi(y)
\]
\[
= (\phi^+(x) + \phi^-(x))(\phi^+(y) + \phi^-(y))
\]
\[
= \phi^+(x)\phi^+(y) + \phi^-(x)\phi^+(y) + \phi^+(x)\phi^-(y) + \phi^-(x)\phi^-(y)
\]
\[
= \phi^+(x)\phi^+(y) + \phi^-(x)\phi^+(y) + \phi^-(y)\phi^+(x) + [\phi^+(x),\phi^-(y)] + \phi^-(x)\phi^-(y)
\]
\[
= :\phi(x)\phi(y): + [\phi^+(x),\phi^-(y)]
\]
\[
= :\phi(x)\phi(y): + D(x-y)
\]

where we have achieved normal ordering at the price of an extra term \([\phi^+(x),\phi^-(y)] = D(x-y)\), which is the propagator of Eq. (5.74).\(^{18}\) Similarly for \(y^0 > x^0\),

\[
T\{\phi(x)\phi(y)\} = :\phi(x)\phi(y): + [\phi^+(y),\phi^-(x)]
\]
\[
= :\phi(x)\phi(y): + D(y-x)
\]

Putting both together, we finally obtain

\[
T\{\phi(x)\phi(y)\} = :\phi(x)\phi(y): + \Delta_F(x-y) \quad (6.27)
\]

where \(\Delta_F(x-y)\) is the **Feynman propagator!** As derived in Sec. 5.8, it has the integral representation

\[
\Delta_F(x-y) = \int \frac{d^3p}{(2\pi)^3} \frac{i}{p^2 - m^2 + i\varepsilon} e^{-ip \cdot (x-y)}
\]

An alternative derivation proceeds as follows: we have shown above, that the difference between normal ordering and time ordering \(\phi(x)\phi(y)\) is just a \(c\)-number.\(^{19}\) So we can equally switch it between

\[
\phi(x)\phi(y)
\]

\[
\phi(x)\phi(y)
\]

Note, that these are either free fields in the Heisenberg picture, or interaction picture fields \(\phi_I(x)\) for \(H = H_0 + H_1\).

\(^{17}\) It is due to the sometimes used notion of **positive** energies in the time evolution of \(\phi^+ \sim e^{-iE_p t}\) and the **negative** energies in the time evolution of \(\phi^- \sim e^{-i(-E_p) t}\).

\(^{18}\) Recall,

\[
D(x-y) = \langle 0 | \phi(x)\phi(y) | 0 \rangle
\]
\[
= \langle 0 | :\phi(x)\phi(y): + [\phi^+(x),\phi^-(y)] | 0 \rangle
\]
\[
= \langle 0 | [\phi^+(x),\phi^-(y)] | 0 \rangle
\]
\[
= [\phi^+(x),\phi^-(y)]
\]

because in the last line, the commutator is a \(c\)-number.

\(^{19}\) The operator structure is lost once we use the commutator for \(a\) and \(a^\dagger\).
the vacuum
\[ \langle \phi(x)\phi(y) \rangle = \langle 0|\phi(x)\phi(y)|0 \rangle \]
\[ = \langle 0|T\{\phi(x)\phi(y)\}|0 \rangle - \langle 0|\phi(x)\phi(y) : |0 \rangle \]
\[ = \langle 0|T\{\phi(x)\phi(y)\}|0 \rangle \]
\[ = \Delta_F(x - y) \]

For **charged fields**, it is straight-forward to show that the propagator is
\[ \langle \psi(x)\psi(y) \rangle = \langle 0|\psi(x)\psi(y)|0 \rangle = \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - M^2 + i\varepsilon} e^{-ip(x-y)} \]
(6.28)

Note, that the only difference is the mass \( M \) of \( \psi \). All other contractions vanish
\[ \langle \psi(x)\psi(y) \rangle = \langle 0|\psi(x)\psi(y)|0 \rangle = 0 \]
as they should. The last equation is easy to show since \( \psi \sim c^\dagger + b \) and \( \psi \) therefore only creates \( c \)-type particles and annihilates \( b \)-type particles, ergo \( \langle 0|T\{\psi(x)\psi(y)\}|0 \rangle = 0 \).

We can now state **Wick’s theorem**. For any collection of fields \( \phi_1 \equiv \phi_{a_1}(x_1), \phi_2 \equiv \phi_{a_2}(x_2) \ldots \) the \( T \)-product of the fields has the expansion
\[ T\{\phi_1 \ldots \phi_n\} = : \phi_1 \ldots \phi_n + \text{all possible contractions} : \]

For \( n = 2 \), it is identical to Eq. (6.26). The term "all possible contractions" means, there will be one term for each possible contraction of the \( n \) fields in pairs. The proof that shows this for all \( n \) is by induction and so not very illuminating. With this theorem, we obtain thus for \( n = 4 \),
\[ T\{\phi_1\phi_2\phi_3\phi_4\} = : \phi_1\phi_2\phi_3\phi_4 + \phi_1\phi_2\phi_3\phi_4 + \phi_1\phi_2\phi_3\phi_4 + \phi_1\phi_2\phi_3\phi_4 + \phi_1\phi_2\phi_3\phi_4 \]
\[ + \phi_1\phi_2\phi_3\phi_4 + \phi_1\phi_2\phi_3\phi_4 + \phi_1\phi_2\phi_3\phi_4 \]
\[ + \phi_1\phi_2\phi_3\phi_4 + \phi_1\phi_2\phi_3\phi_4 + \phi_1\phi_2\phi_3\phi_4 : \]

where a contraction across operators which are not adjacent, still gives a factor of \( \Delta_F \), e.g.
\[ : \phi_1\phi_2\phi_3\phi_4 : = \Delta_F(x_1 - x_3) : \phi_2\phi_4 : \]

6.4.1 **Example: nucleon scattering** \( \psi\psi \rightarrow \psi\psi \)

We are scattering \( \psi\psi \rightarrow \psi\psi \) and therefore have the in- and outgoing states
\[ |i\rangle = |p_1p_2\rangle = \sqrt{2E_{p_1}} \sqrt{2E_{p_2}} b_{p_1}^\dagger b_{p_2}^\dagger |0\rangle \]
\[ |f\rangle = |q_1q_2\rangle = \sqrt{2E_{q_1}} \sqrt{2E_{q_2}} b_{q_1}^\dagger b_{q_2}^\dagger |0\rangle \]
(6.29) (6.30)

**Proof:** Suppose it is true for \( \phi_2 \ldots \phi_n \) and we now add \( \phi_1 \). We choose \( r_k^1 > r_k^0 \) for all \( 2 \leq k \leq n \). We move \( \phi_1 \) out to the left of the time ordered product
\[ T\{\phi_1\phi_2 \ldots \phi_n\} \]
\[ = (\phi_1 \phi_2 \ldots \phi_n) \cdot (\phi_2 \ldots \phi_n + \text{contractions}) \]
The \( \phi_k^1 \sim a^\dagger \) is already normal ordered. To have the RHS as a normal ordered product, we need to commute \( \phi_k^1 \) past the \( \phi_k^0 \) operators. Each time \( \phi_k^1 \) moves past one \( \phi_k^0 \), we pick up a factor
\[ : \phi_1 \phi_k : = \Delta_F(x_1 - x_k) \]
Try it!
This proves Wick’s theorem by induction.
We now want to obtain the expansion of \( \langle f | S|i \rangle \). In fact, we are interested in \( \langle f | S - \mathbb{1}|i \rangle \), since we do not care about amplitudes where nothing happens. At second order in \( g \), we get for \( \langle f | S - \mathbb{1}|i \rangle \)

\[
\left( -i g \right)^2 \int \frac{d^4 x_1 d^4 x_2}{(2\pi)^4} \psi^\dagger(x_1) \psi(x_1) \phi(x_1) \psi^\dagger(x_2) \psi(x_2) \phi(x_2) \quad (6.31)
\]

Wick’s theorem can relate this to a sum of normal-ordered products. The interesting piece for us is

\[
\psi^\dagger(x_1) \psi(x_1) \psi^\dagger(x_2) \psi(x_2) : \phi(x_1) \phi(x_2) \quad (6.32)
\]

which contributes to \( \psi \psi \rightarrow \psi \psi \) scattering, because the two \( \psi \) fields annihilate the initial nucleons and the two \( \psi^\dagger \)'s creates the outgoing nucleons (or \( \psi \)-particles). Any other contraction will give a zero result. We obtain with Eq. (5.68)

\[
\langle q_1 q_2 | : \psi^\dagger(x_1) \psi(x_1) \psi^\dagger(x_2) \psi(x_2) : | p_1 p_2 \rangle = \langle q_1 q_2 | \psi^\dagger(x_1) \psi(x_1) \psi^\dagger(x_2) | 0 \rangle \langle 0 | \psi(x_1) \psi(x_2) | p_1 p_2 \rangle
\]

\[
= \left\{ e^{i q_1 \cdot x_1 + i q_2 \cdot x_2} + e^{i q_1 \cdot x_1 - i q_2 \cdot x_2} + e^{-i p_1 \cdot x_2 - i p_2 \cdot x_1} + e^{-i p_1 \cdot x_2 + i p_2 \cdot x_1} + e^{i x_1 \cdot (q_1 - p_1) + i x_2 \cdot (q_2 - p_2)} + (x_1 \leftrightarrow x_2) \right\}
\]

Let us plug this into Eq. (6.31) to obtain

\[
\left( -i g \right)^2 \int \frac{d^4 x_1 d^4 x_2}{(2\pi)^4} \left\{ \ldots + (x_1 \leftrightarrow x_2) \right\} \int \frac{d^4 k}{(2\pi)^4} \frac{i}{k^2 - m^2 + i\varepsilon} e^{-ik \cdot (x_1 - x_2)}
\]

where the \( \{ \ldots \} \) corresponds to the sum of exponentials above resulting from the matrix element of the normal-ordered operator. Since the propagator is symmetric under \( x_1 \leftrightarrow x_2 \) we can drop this in exchange for the \( 1/2 \) in front. The \( x_i \) integrals result in \( \delta \)-functions, e.g. for the first term we get for the exponentials

\[
\int d^4 x_1 d^4 x_2 e^{i x_1 \cdot (q_1 - p_1) + i x_2 \cdot (q_2 - p_2) - ik \cdot (x_1 - x_2)} = (2\pi)^8 \delta^{(4)}(q_1 - p_1 - k) \delta^{(4)}(q_2 - p_2 + k)
\]

and we find

\[
i(-ig)^2 \int \frac{d^4 k}{(2\pi)^4} \frac{(2\pi)^8}{k^2 - m^2 + i\varepsilon} \left\{ \delta^{(4)}(q_1 - p_1 - k) \delta^{(4)}(q_2 - p_2 + k) + \delta^{(4)}(q_1 - p_2 + k) \delta^{(4)}(q_2 - p_1 - k) \right\}
\]

We finally obtain

\[
i(-ig)^2 \left[ \frac{1}{(p_1 - q_1)^2 - m^2} + \frac{1}{(p_2 - q_2)^2 - m^2} \right] (2\pi)^4 \delta^{(4)}(p_1 + p_2 - q_1 - q_2)
\]

Notice that performing the final integral over \( \delta \)-functions leaves us with a factor of

\[
(2\pi)^4 \delta^{(4)}(p_1 + p_2 - q_1 - q_2)
\]

which just enforces 4-momentum conservation for the scattering process. Since we usually have Poincaré invariant Lagrangians, it

---

20 The two nucleons in this case just continue to fly without any interaction as free particles.

21 Poincaré invariant = space-time translation (and Lorentz) invariant. For the \( \delta^{(4)} \) function the space-time translation invariance is the crucial part, since it guarantees total 4-momentum conservation.
is customary to define the invariant **Feynman amplitude** $A_{fi}$ by

$$\langle f | S - 1 | i \rangle = i A_{fi} (2\pi)^4 \delta^{(4)}(p_f - p_i)$$

where the factor $i$ is by convention, since it reproduces the phase conventions for scattering in non-relativistic quantum mechanics.

Further, $p_\mu^f = \sum_n p_\mu^{f,n}$ and $p_\mu^i = \sum_n p_\mu^{i,n}$. We can ignore the $i\varepsilon$, because the denominator is never zero: in the center of mass frame, we can write the momenta as

$$p_1 = (\sqrt{p^2 + M^2}, p, 0, 0) \quad (6.33)$$
$$p_2 = (\sqrt{p^2 + M^2}, -p, 0, 0) \quad (6.34)$$
$$q_1 = (\sqrt{p^2 + M^2}, p \cos \theta, p \sin \theta, 0) \quad (6.35)$$
$$q_1 = (\sqrt{p^2 + M^2}, -p \cos \theta, -p \sin \theta, 0) \quad (6.36)$$

which immediately gives

$$(p_1 - q_1)^2 = -2p^2(1 - \cos \theta), \quad (p_1 - q_2)^2 = -2p^2(1 + \cos \theta)$$

and so

$$A = g^2 \left[ \frac{1}{2p^2(1 - \cos \theta) + m^2} + \frac{1}{2p^2(1 + \cos \theta) + m^2} \right] \quad (6.37)$$

Note that the two terms are required because of Bose statistics. Scattering into two identical particles at an angle $\theta$ is indistinguishable from scattering at an angle $\theta - \pi$, and so the probability must be symmetrical under the interchange of the two processes. Since these particles are bosons, the amplitude must also be symmetric.

This was a bit tedious. There is a much simpler way which we will discuss in a little bit using Feynman diagrams.

### 6.5 Diagrammatic Perturbation theory: Feynman Diagrams

Our final result for $\psi \psi \rightarrow \psi \psi$ scattering in Eq. (6.37) was remarkably simple, even though the intermediate steps were a bit messy.

This motivates a diagrammatic short-hand called **Feynman diagrams**. These are pictures of the fields and contractions which we need to evaluate to give the matrix element. At the $n^{th}$ order in perturbation theory $H_I$ will act $n$ times and so a Feynman diagram will contain $n$ interaction vertices. For our toy model the interaction vertex looks like in Fig. 6.3.

We need to distinguish $\psi$ from $\psi^\dagger$ and we will therefore draw an arrow on the line.

Next, **contractions** are the lines joining different vertices. Every time there is a contraction, we join the lines of the contracted fields. The example of Eq. (6.32) corresponds to the diagram Fig. 6.4.

Arrows will line up because $\psi(x) \psi(y) = \psi^\dagger(x) \psi^\dagger(y) = 0$ and the arrows **indicate the flow of the $U(1)$ charge**. Now, any fields
which are left uncontracted must either annihilate particles from the incoming state or create particles from the outgoing state. If there are different ways of doing this, they correspond to indistinguishable processes, so we must add the corresponding amplitudes: we write down a separate Feynman diagram for each distinct labeling of the external legs. For \( \psi \psi \rightarrow \psi \psi \) scattering, this results in two Feynman diagrams.

We will revisit them below.

6.5.1 Feynman rules

1. For each particle in \(|i\rangle\) and \(|f\rangle\), draw an external line. Assign a directed momentum \(p_i\) for each line, and add an arrow to denote its charge.

2. Join the external lines together with vertices as in Fig. 6.3.

3. At each vertex, write a factor of

\[
(-ig)(2\pi)^4\delta^{(4)}(\sum k_i)
\]

with \(\sum k_i\) is the sum of all momenta flowing into the vertex.

4. For each internal \(\phi\) line, we write a factor of

\[
\int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2 - m^2 + i\varepsilon}
\]

and for each internal \(\psi\) line, we write

\[
\int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2 - M^2 + i\varepsilon}
\]

5. Divide the final result by the overall energy-momentum conserving \(\delta\)-function, \((2\pi)^4\delta^{(4)}(p_f - p_i)\).

In fact, we can simplify things further and get rid of the trivial delta-functions: we impose energy-momentum conservation on the momenta flowing into each vertex and so we change

3.’ At each vertex, write a factor

\[
(-ig)
\]

4.’ Each contracted internal line represents a propagator (or \(M^2\) for internal \(\psi\) lines)

\[
\frac{i}{k^2 - m^2 + i\varepsilon}
\]

This is ok for graphs like the ones we have been considering. However, there are also diagrams with closed loops where energy-momentum conservation at the vertices does not fully fix all the internal momenta! For example the matrix element obtained from

![Figure 6.5](image1.png)

Figure 6.5: Feynman diagrams contributing at order \(g^2\) to \(\psi \psi \rightarrow \psi \psi\) scattering, see Eq. (6.32). We show the explicit momentum assignment for the external and internal lines. There are two distinct options.

![Figure 6.6](image2.png)

Figure 6.6: Diagrammatic representation of Eq. (6.38) where the blue arrows show the momentum assignments.
This means, we need to keep the factor

$$\int \frac{d^4k}{(2\pi)^4}$$

Similarly for the fully contracted term

$$\langle 0 | : \bar{\psi}(x_1)\psi(x_2)\bar{\psi}(x_1)\psi(x_2)\phi(x_1)\phi(x_2) : | 0 \rangle \quad (6.39)$$

neither $p$ nor $k$ is constrained (not fixed by external momenta) and we must integrate over both momenta. Thus we must add a rule that we integrate with $\int \frac{d^4k}{(2\pi)^4}$ over each internal line whose momentum $k$ is unconstrained.
Let us summarize our simplified and final **Feynman rules** to calculate $iA_{f_i}$ which we defined as $\langle f | S - 1 | i \rangle = iA_{f_i} (2\pi)^4 \delta^{(4)}(p_f - p_i)$

1. For each particle in $|i\rangle$ and $|f\rangle$, draw an external line. Assign a directed momentum $p_i$ for each line, and add an **arrow** to denote the flow of charge.

2. Join the external lines together with vertices as in Fig. 6.3.

3. At each vertex, write a factor of

   $$(-ig)$$

   and impose 4-momentum conservation.

4. For each internal $\phi$ line, we write a factor of

   $$\frac{i}{k^2 - m^2 + i\varepsilon}$$

   and for each internal $\psi$ line, we write

   $$\frac{i}{k^2 - M^2 + i\varepsilon}$$

5. For each loop with momentum $k$ unconstrained write a factor of

   $$\int \frac{d^4k}{(2\pi)^4}$$
6.5.2 Revisiting $\psi\psi \rightarrow \psi\psi$ scattering

Let us apply the above Feynman rules to compute the amplitude of $\psi\psi \rightarrow \psi\psi$ scattering at order $g^2$. The two diagrams contributing to this process at this order are and we obtain using the Feynman rules

\[ iA = (-ig)^2 \left[ \frac{i}{(p_1 - q_1)^2 - m^2} + \frac{i}{(p_1 - q_2)^2 - m^2} \right] \]  

(6.40)

The interpretation of these diagrams is the following: the nucleons $\psi$ exchange a meson $\phi$ which has momentum $k = p_1 - q_1 = p_2 - q_2$. This meson does not satisfy the usual energy momentum relation, since as we saw above $k^2 = -2p^2(1 - \cos \theta) < 0$ and therefore in particular $k^2 \neq m^2$. The meson $\phi$ is called a virtual particle and which is off-shell.\(^{25}\) Heuristically we can say, it cannot live long enough for its energy to be measured to great accuracy.

6.6 More scattering processes in the $\bar{\psi}\psi\phi$ theory

Let us follow Feynman’s motto and let us ’shut up and calculate’ a couple of sample processes.

6.6.1 Nucleon anti-nucleon scattering: $\psi(p_1) + \bar{\psi}(p_2) \rightarrow \psi(q_1) + \bar{\psi}(q_2)$

There are two diagrams contributing to this process. Note the opposite flow of the arrows on the Feynman-lines indicating the difference between a nucleon and an anti-nucleon. We obtain,

\[ iA = (-ig)^2 \left[ \frac{i}{(p_1 - q_1)^2 - m^2} + \frac{i}{(p_1 + p_2)^2 - m^2} \right] \]  

(6.41)

Note, since the diagrams are simply a shorthand for matrix elements of operators in the Wick expansion, the orientation of the lines inside the graphs have absolutely no significance. In the second contribution, the virtual meson can go on-shell for certain values of}

\(^{25}\) $k^2 = m^2$ would be on-shell.
the external momenta. In the center of mass frame $p_1 = -p_2$ and
the 2nd denominator is
\[
\frac{1}{(p_1 + p_2)^2 - m^2} = \frac{1}{4(M^2 + p_1^2) - m^2}
\]
which can lead to a divergence if $m > 2M$ for certain values of $p_1$.
This will be fixed if one adds higher order corrections to the Feynman
graph in the form of an imaginary contribution (the width) and
we will find a bump if we scan the momenta $p_1$ in an experiment.

6.6.2 Meson-meson scattering $\phi\phi \rightarrow \phi\phi$

For $\phi\phi \rightarrow \phi\phi$, the simplest diagram appears only at $g^4$. One of the
momenta remains unconstrained

\[
\begin{align*}
&\begin{array}{c}
p_1 \uparrow \quad \downarrow \quad k - q_1 \\
&\end{array} \\
&\begin{array}{c}
k - q_1 + p_1 \\
&\end{array} \\
&\begin{array}{c}
p_2 \uparrow \quad \downarrow \quad k + q_2 \\
&\end{array}
\end{align*}
\]

and we need to integrate over it,
\[
iA = \int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2 - M^2 + i\varepsilon} \frac{i}{(k - q_1)^2 - M^2 + i\varepsilon} \\
\times \frac{i}{(k - q_1 + p_1)^2 - M^2 + i\varepsilon} \frac{i}{(k + q_2)^2 - M^2 + i\varepsilon}
\]
This is a one-loop diagram and the integration of $k^\mu$ can require
some additional insights as diagrams can be divergent. This was a
serious problem in the early days of quantum field theory but we
now understand it along the lines of the discussion in Sec. 5.1.2. In
the example at hand we are fine though, since the integration is
convergent, since $iA \sim \int d^4k/k^8$ for very large $k^\mu$.

6.6.3 Symmetry factors

Are there combinatoric considerations in associating operators with
Feynman diagrams? Write the expansion of Eq. (6.20)
\[
S = T \exp \left( -i \int d^4x \mathcal{H}_I(x) \right)
\]
in powers of the interaction
\[
S = \sum_{n=0}^{\infty} S^{(n)}
\]
\[
= \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int d^4x_1 \ldots d^4x_n \mathcal{T} \{ \mathcal{H}_I(x_1) \ldots \mathcal{H}_I(x_n) \}
\]
Generally, we might expect the factor $1/n!$ to get canceled when we carry out the Wick expansion, because we can permute the vertices to find a different contraction that gives the same normal ordered operator. But this counting can change if the diagrams have symmetries, e.g. permutations of the vertices that do not give rise to a new contraction!

Let us consider three Klein-Gordon scalar fields $\phi_1, \phi_2, \phi_3$ with masses $m_1, m_2, m_3$ and an interaction

$$H_I = \lambda \phi_1(x)\phi_2(x)\phi_3(x)$$

and discuss the diagram associated with the contraction

$$\left(\phi_1\phi_2\phi_3\right)\left(\phi_1\phi_2\phi_3\right)\left(\phi_1\phi_2\phi_3\right)$$  \hspace{1cm} (6.44)

There are $3!$ contractions associated with this diagram, and the $1/3!$ in front of the $S(3)$ gets completely canceled.

We now need to add to our list of rules associating operators with diagrams, we may need to add one more: Divide by the symmetry factor of the graph. The symmetry factor is the number of permutations of the vertices that leave the contraction unchanged.

In the $\lambda \phi_1(x)\phi_2(x)\phi_3(x)$-theory, this consideration will give the complete symmetry factor. In general, another contribution to the symmetry factor of a graph occurs if there are identical fields at a single vertex.

Let us illustrate this with in $\lambda \phi^3(x)$-theory, where we consider a single real KG-scalar field of mass $m$ with the interaction

$$H_I = \frac{\lambda}{3!} \phi^3(x)$$

The factor $1/3!$ is included since there are usually $3!$ different ways of contracting the fields of a vertex with the fields of neighboring vertices, and so the $1/3!$ gets canceled.

However, the $1/3!$ does not always get completely canceled because permutations of the lines might not give a new contraction. We must therefore also include in the symmetry factor of a graph the number of permutations of lines that leave the contraction unchanged.
You can convince yourself, that we can put these considerations in a simple formula for the symmetry factor $S(G)$ of a graph $G$

$$S(G) = \frac{n!(\eta)^n}{r}$$

where $n$ is the number of vertices, $\eta$ comes from the coupling constant and is $\eta = 4!$ in $\frac{\lambda}{\eta^4} \phi^4$ theory and $\eta = 3!$ in $\frac{\lambda}{\eta^3} \phi^3$ theory, and $r$ is the multiplicity of the diagram. The multiplicity $r$ is determined by labeling all vertices with three (or four) lines emerging. We assume all these lines to be distinguishable. The multiplicity $r$ is then the total number of ways to connect external points and vertices to form the diagram.\(^{26}\)

To get the correct overall constant for a diagram, we divide by its symmetry factor.

**Examples:** We will use the $\frac{\lambda}{3!} \phi^3(x)$ and $\frac{\lambda}{4!} \phi^4(x)$ theory

- **No symmetry since all three external lines are distinct:**

  \[
  S(G) = 1
  \]

  which comes about from

  \[
  \langle q_1 q_2 | : \frac{\lambda}{3!} \phi(x) \phi(x) \phi(x) : | p_3 \rangle = \frac{\lambda}{3!} 3! \langle q_1 q_2 | : \phi(x) \phi(x) \phi(x) : | p_3 \rangle
  \]

  we can contract each of the three field operators with each of the external fields. Let us check by calculating

  \[
  S(G) = \frac{1!(3!)^1}{3!} = 1 \quad (6.45)
  \]

  where $\eta = 3!$ and $r = 3!$ since there are 3! ways to connect the external points to the 3 lines emerging from the vertex.

- **This is a 2nd order graph in $\frac{\lambda}{3!} \phi^3$ theory:**

  \[
  S(G) = \frac{2!(3!)^2}{(3 \cdot 2)^2} = 1
  \]

- **This is a 2nd order graph in $\frac{\lambda}{4!} \phi^4$ theory:**

  \[
  S(G) = \frac{2!(4!)^2}{8 \cdot 3 \cdot 4 \cdot 3 \cdot 2} = 2
  \]

  with $r = 8 \cdot 3 \cdot 4 \cdot 3 \cdot 2$ because we can connect the first external point to 8 different lines emerging from the two vertices, after
that we can use the remaining 3 lines and the 4 · 3 lines from the other vertex. The internal propagator is then fixed apart from a two-fold ambiguity.

- This is a 2nd order graph in $\frac{1}{\pi^4}\phi^4$ with two loops

$$S(G) = \frac{2!(4!)^2}{8 \cdot 4 \cdot 3 \cdot 2} = 6$$

with $r = 8 \cdot 4 \cdot 3 \cdot 3$ because there are $8 \cdot 4$ ways to connect to the external points and once we choose one of the internal lines we have first 3 options, then 2 to connect it to the other vertex.

- A 4th order $\frac{1}{\pi^3}\phi^3$ graph:

$$S(G) = \frac{4!(3!)^4}{12 \cdot 2 \cdot 9 \cdot 2 \cdot 6 \cdot 2 \cdot 3 \cdot 2} = 1$$

Most people never need to evaluate a diagram with a symmetry factor greater than 2, so you should not worry too much about this technicality.

6.7 Potentials

Let us look at the non-relativistic limit of our scattering amplitude with the goal to treat it as a potential in non-relativistic quantum mechanics.

6.7.1 The Yukawa Potential

We first re-examine the problem of a classical Klein-Gordon field in the presence a point source. We will just repeat the steps of Sec. 4.3.2 where we now have to solve the static Klein-Gordon equation

$$(\Box + m^2)\phi(x) = \delta^3(x)$$

we can invert this and simplify since $\delta^3(x)$ is time-independent

$$\phi(x) = \frac{1}{-\nabla^2 + m^2} \delta^3(x)$$

We solve again using the Fourier transformation of the $\delta$-function

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{-\nabla^2 + m^2} e^{ix \cdot p}$$

$$= \int \frac{d^3p}{(2\pi)^3} \frac{1}{p^2 + m^2} e^{ix \cdot p}$$

$$= \frac{1}{4\pi^2} \frac{1}{iv} \int_0^\infty dp \frac{e^{ipr} - e^{-ipr}}{p^2 + m^2}$$

$$= \frac{1}{2\pi r} \int_{-\infty}^{\infty} dp \frac{e^{ipr}}{p^2 + m^2}$$
where we have skipped in the 2nd and 3rd line the steps which we have derived carefully in Sec. 4.3.2. We close the contour in the upper half plane \( p \to +i\infty \) where we pick up the pole at \( p = +im \), which finally gives

\[
\phi(x) = \frac{1}{4\pi r} e^{-mr}
\]

(6.46)

If you compare this to the \( \sim 1/r \) of Coulomb’s law Eq. (4.95) you see that the field now dies off exponentially at distances \( > 1/m \), which is the Compton wavelength of the meson.

Can we now understand the profile of the \( \phi \) field as a force between \( \psi \) particles? Recall, that we treat the electro-static potential sourced by a \( \delta \)-function as the potential energy for another charged (test) particle moving in this background.

Is there a classical limit of the scalar Yukawa theory with the \( \psi \) particles as \( \delta \)-function sources for \( \phi \), creating the Yukawa potential? The answer is yes, at least in the limit \( M \gg m \), that is, when the nucleons are much heavier than the mesons and act as quasi static sources.

Recall the Born approximation\(^{27}\) in non-relativistic quantum mechanics: at first order in perturbation theory, we find the amplitude to scatter an incoming state with momentum \( p \) of a potential \( U(r) \) into an outgoing state with momentum \( q \)

\[
\mathcal{A}_{QM}(p \to q) = \langle q | U(r) | p \rangle = -i \int d^3r \ U(r) e^{-i(q-p) \cdot r}
\]

We will now compare the non-relativistic potential scattering with that from the first diagram in Eq. (6.40) in the center of mass frame

\[
iA = -g^2 \frac{i}{(p_1 - q_1)^2 - m^2} = g^2 \frac{i}{|p_1 - q_1|^2 + m^2}
\]

(6.47)

where we have used the fact that in the center of mass frame the energies of the incoming and scattered particles are the same, see Eq. (6.33).

We need to account for the different normalizations of the relativistic and non-relativistic amplitudes and divide the result by \( (2M)^2 \) to account for the difference between relativistic and non-relativistic normalization of states, see Sec. 5.4. We define the dimensionless quantity

\[
\lambda \equiv \frac{g}{2M}
\]

and obtain

\[
\int d^3r \ U(r) e^{-i(q_1 - p_1) \cdot r} = -\frac{\lambda^2}{|p_1 - q_1|^2 + m^2}
\]

Inverting the Fourier-transform gives

\[
U(r) = -\lambda^2 \int \frac{d^3p}{(2\pi)^3} e^{-ip \cdot r} \frac{1}{p^2 + m^2}
\]

\(^{27}\) See e.g. Cohen-Tannoudji, Quantum Mechanics, Vol. II, Chapter VIII, especially section B.4
This is exactly the integral, that we already calculated in the beginning of the section and we find

\[ U(r) = -\frac{\lambda^2}{4\pi r} e^{-mr} \]

This is the **Yukawa potential**, where the minus sign tells us that it is attractive. Yukawa made this potential\(^{28}\) the basis for his theory of the nuclear force and worked backwards from the range of the force (about 1 fm) to predict the mass (about 200 MeV) of the required meson, the pion.

We can directly from the scattering amplitudes for \( \psi\bar{\psi} \rightarrow \psi\bar{\psi} \), \( \psi\bar{\psi} \rightarrow \psi\bar{\psi} \), and \( \psi\bar{\psi} \rightarrow \psi\bar{\psi} \) that the **sign** of the spin-0 generated Yukawa term is always positive – it leads to a universally attractive potential. Compare this to the Coulomb potential which is mediated by the exchange of a spin-1 boson, which can be either attractive and repulsive. Gravity, which is mediated by a spin-2 field is again universally attractive.

Notice that quantum field theory has given us an entirely new view of forces between particles. Rather than being a fundamental concept, the force arises from the virtual exchange of other particles, in this case the meson.

### 6.8 How to calculate observables

We are now able to calculate amplitudes for a set of processes by evaluating Feynman diagrams,

\[ \langle f | S - 1 | i \rangle = i A_{fi} (2\pi)^4 \delta^{(4)}(p_F - p_I) \]  \hspace{1cm} (6.48)

but we have yet to calculate a probability, which would be measurable. For this, we need to square the amplitudes and sum over all observed final states. But it looks like we are in trouble here, since

\[ |\langle f | S - 1 | i \rangle|^2 = |A_{fi}|^2 \left[ (2\pi)^4 \delta^{(4)}(p_F - p_I) \right]^2 \]  \hspace{1cm} (6.49)

Squaring a \( \delta \)-function makes no sense. What happened?

As we will see, the problem is that we are not working with square-integrable states. Instead, our states are normalized to \( \delta^{(3)} \)-functions, since they are plane waves which exist at every point in space-time. Thus, the scattering process is occurring at all points in space for all time. No suprise that we got divergent non-sense.

#### 6.8.1 Fermi’s Golden Rule

Let us derive the familiar Fermi’s golden rule from Dyson’s formula of Eq. (6.15). Consider two energy eigenstates \( |n\rangle \) and \( |m\rangle \) of \( H_0 \), with \( E_n \neq E_m \). We find in leading order in the perturbation

Recall from Eq. (6.5), that the interaction Hamiltonian in the interaction picture is

\[ H_I = (H_{int})_I = e^{iH_{0}t}(H_{int})e^{-iH_{0}t} \]
golden rule for the transition rate $\Gamma$ amplitude, since $P$

Squaring this, we would run into the same problem as for the Feynman-amplitude, since $P_{n\to m}(t) = |\langle m|H_{\text{int}}|n \rangle|^2 (2\pi)^2 \delta(\omega)^2$, a non-sensical

we take the limit for states around the same energy $E$. We finally obtain a constant transition probability per unit time $\rho$

$\Delta E \approx \frac{2\pi}{t}$

As we take $t \to \infty$, the function in brackets starts to approach a $\delta$-function. We get the normalization by

\[ \int_{-\infty}^{\infty} d\omega \left[ \frac{4\sin^2(\omega t/2)}{\omega^2} \right] = 2\pi t \]

and therefore

\[ 4 \frac{\sin^2(\omega t/2)}{\omega^2} \to 2\pi t \delta(\omega), \quad \text{for } t \to \infty \]

We consider now a transition to a cluster of states with density $\rho(E)$. In the limit $t \to \infty$, we get

\[ P_{n\to m}(t) = \int dE_m \rho(E_m) |\langle m|H_{\text{int}}|n \rangle|^2 \frac{4\sin^2(\omega t/2)}{\omega^2} \]

\[ \to 2\pi t |\langle m|H_{\text{int}}|n \rangle|^2 \rho(E_m) \]

We finally obtain a constant transition probability per unit time for states around the same energy $E_n \sim E_m = E$. This is Fermi’s golden rule for the transition rate $\Gamma_{n\to m} = \hat{P}_{n\to m}$

$$\hat{P}_{n\to m} = 2\pi |\langle m|H_{\text{int}}|n \rangle|^2 \rho(E) \quad (6.50)$$

This was a useful exercise, since it showed us how to carefully take the limit $t \to \infty$. If we chose to compute the amplitude for the state $|n \rangle$ at $t \to -\infty$ to transition to $|m \rangle$ at $t \to +\infty$ we would get the amplitude

\[ -i\langle m|\int_{-\infty}^{\infty} dt' H_I(t')|n \rangle = -i\langle m|H_{\text{int}}|n \rangle 2\pi \delta(\omega) \]

Squaring this, we would run into the same problem as for the Feynman-amplitude, since $P_{n\to m}(t) = |\langle m|H_{\text{int}}|n \rangle|^2 (2\pi)^2 \delta(\omega)^2$, a non-sensical

Note however, according to Wikipedia: “Although named after Enrico Fermi, most of the work leading to the Golden Rule is due to Paul Dirac who formulated 20 years earlier a virtually identical equation, including the three components of a constant, the matrix element of the perturbation and an energy difference. It was given this name because, on account of its importance, Fermi dubbed it ‘Golden Rule No. 2.’ ” This seems to show how to avoid the Matthew principle, but you need to be yourself famous enough. ‘Golden Rule No. 1.’ states the second-order contributions to the transition probability.
infinity. Comparing to the result above, we realize now that the extra infinity is coming because $P_{n\rightarrow m}$ is the probability for the transition to happen in infinite time $t \rightarrow \infty$. We can write these $\delta$-functions as

$$(2\pi)^2 \delta(\omega)^2 = (2\pi) \delta(\omega) T$$

where $T$ signals $t \rightarrow \infty$. We can now divide by $T$, to derive the probability per unit time as

$$\dot{P}_{n\rightarrow m} = 2\pi |\langle m |H_{\text{int}} | n \rangle|^2 \delta(\omega)$$

which after integrating over the density of final states gives us Fermi’s Golden rule again. We will now interpret the squares of the $\delta(4)$-functions in Eq. (6.49) as space-time volume factors.

### 6.8.2 Decay Rates

We now want to obtain the probability for a single particle $|i\rangle$ of momentum $p_I$ to decay into some number of particles $|f\rangle$ with momentum $p_F = \sum_l p_{l}^f$. This probability is given by

$$P = |\langle f |S| i \rangle|^2 \frac{\delta^{(3)}(0)}{\langle f |f \rangle \langle i |i \rangle}$$

We have to be careful with the normalization since we work with states obeying the relativistic normalization convention, see Sec. 5.4. In particular, with Eq. (5.50), we obtain

$$\langle i |i \rangle = (2\pi)^3 2E_p \delta^{(3)}(0) = 2E_{p_i} V$$

and

$$\langle f |f \rangle = \prod_{\text{final states}} 2E_{p_l} V$$

Let us place our initial particle at rest $p_I = 0$ and $E_{p_i} = m_i$, we obtain for the probability for the decay

$$P = |\langle f |S| i \rangle|^2 \times \frac{1}{\langle f |f \rangle \langle i |i \rangle}$$

$$= |A_{fi}|^2 (2\pi)^4 \delta^{(4)}(p_I - p_F) VT \times \frac{1}{2mV} \prod_{\text{final states}} \frac{1}{2E_{p_l} V}$$

where like in the derivation of the Golden Rule we have replaced one of the $\delta^{(4)}$-functions with a space-time factor $(2\pi)^4 \delta^{(4)} = VT$. We have been already computing the amplitude $A_{fi}$. We can now get rid of $T$ by calculating the transition rate, which we get by taking the $T$ derivative (or just dividing by $T$). One worry is still the sum over all final states. We proceed in two steps: first we put our theory in a box with side $L$ assuming periodic boundary conditions which allows us to count states. Then we integrate over all possible momenta of all final states.

We used a similar trick, when we discussed the vacuum energy in Sec. 5.1.2.

Recall the example of meson decay $\phi \rightarrow \bar{\psi}\psi$ of Eq. (6.25)

$$\langle f |S| i \rangle = -ig (2\pi)^4 \delta^{(4)}(p - q_1 - q_2)$$

$$= iA_{fi} (2\pi)^4 \delta^{(4)}(p - q_1 - q_2)$$

and therefore

$$A_{fi} = -g$$

in leading order.
As we discussed in the introduction, in a box of dimension $L^3$ with periodic boundary conditions, the allowed momentum values are of the form

$$k_i = \frac{2\pi n_i}{L}, \quad \text{with } n_i = 0, \pm 1, \pm 2, \ldots, \text{ and } i = x, y, z$$

which implies a density of states $(L/(2\pi))^3$ in momentum space.

In the real world, no experimentalist can measure a cross-section with exact momenta assignments. We can only measure all states about some small region $\Delta k$ in momentum space. In a region of size $\Delta k_x \Delta k_y \Delta k_z$, there are

$$\Delta k_x \Delta k_y \Delta k_z \times \left( \frac{L}{2\pi} \right)^3 = \frac{V}{(2\pi)^3} \Delta k_x \Delta k_y \Delta k_z$$

states.

The number of final one-particle states in an infinitesimal momentum interval is

$$\frac{V}{(2\pi)^3} d^3 k$$

and we find for the density of final states that the factors of $V$ cancel

$$d\Pi = (2\pi)^3 \delta^{(4)}(p_I - p_F) \prod_{\text{final states}} \frac{d^3 p_I}{(2\pi)^3} \frac{1}{2E_{p_I}}$$

(6.53)

and find a Lorentz-invariant quantity, since the $1/(2E_{p_I})$ conspire to give us a Lorentz-invariant measure. This gives us our final expression for the decay probability per unit time $\Gamma = \dot{P}$,

$$\Gamma = \frac{1}{2m} \sum_{\text{final states}} |A_{fi}|^2 d\Pi$$

(6.54)

where the sum is over final states with different numbers and types of particles. After a time $t$, the probability that the particle has not decayed is just $e^{-\Gamma t}$. We call $\Gamma$ the width of the particle. It is the inverse of the half-life $\tau = 1/\Gamma$. If we consider the uncertainty principle, we can see this: the particle exists for a time $\tau$, any measurement of its energy (or mass in the rest-frame) must be uncertain by $\sim 1/\tau = \Gamma$. Therefore, a series of measurements of the particle mass will have a characteristic spread of order $\Gamma$.

Since we will need it later, we will also derive a general expression for the probability per unit time using Eq. (6.52)

$$\dot{P} = |A_{fi}|^2 V \cdot \prod_{\text{initial states}} \frac{1}{2E_{p_I}} \frac{V}{\sqrt{\prod_{\text{initial states}} 2E_{p_I}}} \cdot d\Pi$$

(6.55)

where we have used Eq. (6.53) for the phase space of the final-states.
6.8.3 Example: two-body phase space

For two particles, there are six integrals in $d^3p_1 d^3p_2$ but four of these variables are constrained by 4-vector conservation $\delta^{(4)}(p_1 + p_2 - p_T)$, leaving only two variables to integrate over.

$$d\Pi = \frac{d^3p_1}{(2\pi)^3} \frac{1}{2E_{p_1}} d^3p_2 \frac{1}{(2\pi)^3} \frac{1}{2E_{p_2}} (2\pi)^4 \delta^{(4)}(p_1 + p_2 - p_T)$$

In the center of mass frame, $p_T = 0$ and $E_T \equiv E_T$ which is total energy in the process.

$$d\Pi = \frac{d^3p_1}{(2\pi)^3} \frac{1}{2E_{p_1}} \frac{d^3p_2}{(2\pi)^3} \frac{1}{2E_{p_2}} (2\pi)^3 \delta(E_1 + E_2 - E_T) (2\pi)^3 \delta(p_1 + p_2)$$

$$= \frac{1}{(2\pi)^3} \frac{1}{4E_{p_1} E_{p_2}} (2\pi)^3 \delta(E_1 + E_2 - E_T)$$

We now need to convert the $\delta$-function of energy into a constraint on $p_1$. Using the notation of the margin note, we have $g(p_1) = E_1(p_1) + E_2(p_2) = \sum_1 \sqrt{p_1^2 + m^2}$ and

$$g'(p_1) = \sum_{i=1}^2 \frac{1}{\sqrt{p_1^2 + m^2}} \frac{1}{2}(2p_1) = \frac{p_1}{E_1} + \frac{p_1}{E_2}$$

$$= \frac{p_1(E_1 + E_2)}{E_1 E_2} = \frac{p_1}{E_T} \frac{E_T}{E_1}$$

where we have used $p_1 = -p_2$.

We obtain finally for the two-body phase space in the center of mass frame,

$$d\Pi_2 = \frac{1}{16\pi^2} \frac{p_1}{E_T} d\Omega$$

(6.56)

Here we have assumed that the particles $A$ and $B$ in the final state are distinguishable and we treated the final states $|A(p_1), B(p_2)|$ and $|A(p_2), B(p_1)|$ as distinct. Note, that if the final states are instead identical, we would have double-counted by 2! In general, we need to multiply an $n$-body phase space of $n$ identical particles in the final states by a factor of

$$d\Pi_n \rightarrow \frac{1}{n!} d\Pi_n$$

6.8.4 Examples: particle decays

Let us apply this to an example. We assume that $m > 4M^2$ and the pion to two nucleon decay

$$\phi \rightarrow \bar{\psi}\psi$$

is kinematically allowed. There is only one diagram at leading order in perturbation theory, see Eq. (6.25),

$$iA = -ig$$

We use again that

$$\delta(x) = \sum_i \frac{\delta(x - x_i)}{|g'(x_i)|}$$

where the sum extends over all roots of $g(x)$ which are assumed to be simple. For example

$$\delta(x^2 - \alpha^2) = \frac{1}{2|\alpha|} \left[ \delta(x + \alpha) + \delta(x - \alpha) \right]$$

Figure 6.13: Pion decay into two nucleons. This is kinematically allowed in our toy model if $m > 4M^2$. Note, in nature the pion is the lightest QCD resonance and would not be able to decay to two nucleons.
and the decay width of $\phi$ is therefore

$$
\Gamma(\phi \to \bar{\psi}\psi) = \frac{1}{2m} \int d\Omega |A|^2 \\
= \frac{g^2}{2m} \frac{p_1}{16\pi^2 E_T} \int d\Omega \\
= \frac{g^2}{2m} \frac{p_1}{8\pi m}
$$

where we have used $\int d\Omega = 4\pi$ and $E_T = m$. We adapt the result of Sec. 3.5.2, and find $p_1 = \frac{1}{2} \sqrt{m^2 - 4M^2}$ to obtain

$$
\Gamma(\phi \to \bar{\psi}\psi) = \frac{g^2}{32\pi m} \sqrt{1 - \frac{(2M)^2}{m^2}} \quad (6.57)
$$

### 6.9 Cross Sections

In physical experiments, we often collide two beams of particles or a beam hits a target. From time to time, they hit each other. We then measure the resulting number of particles incident on a detector.

The fraction of the times (per unit flux) they collide is called the cross section and we denote it as $\sigma$. The incoming flux $F$ is defined as the number of incoming particles per area per unit time. The total number of scattering events $N$ per unit time is then given by

$$
N = F \cdot \sigma \quad (6.58)
$$

We can use quantum field theory to calculate $\sigma$. In fact, we can extract even more and we can calculate $d\sigma$, which is the differential cross-section. This $d\sigma$ is related the probability for a given scattering process into a solid angle $(\theta, \phi)$. With $E_1$ and $E_2$ as the energies of the two incoming particles, we obtain with Eq. (6.55)

$$
d\sigma = \frac{\text{differential probability}}{\text{unit time} \cdot \text{unit flux}} \\
= \frac{|A_{fi}|^2}{2E_1 2E_2 V} \cdot \frac{1}{F} \cdot d\Pi 
$$

We now need an expression for the unit flux $F$. Let us assume for simplicity that we are in the center of mass frame of the collision. We have so far considered a single particle per spatial volume $V$ such that the flux is given in terms of the 3-velocities as

$$
F = \frac{|v_1 - v_2|}{V}
$$

We therefore get

$$
d\sigma = \frac{1}{4E_1 E_2 |v_1 - v_2|} \cdot |A_{fi}|^2 \cdot d\Pi \quad (6.59)
$$

You can now use Eq. (6.59) and take your favorite scattering amplitude to compute the probability for particles to end up at various angles.

This is easy to see: consider a beam of particles perpendicular to a plane of area $A$ moving with $v$. If the density of particles is $n$, then after a time $t$, the total number of particles passing through the plane is

$$
N = |v| A \cdot t \cdot n
$$

The flux is therefore $N/(At) = |v| \cdot n$. With our normalization, there is one particle in the box with volume $V$, so $n = 1/V$ and the flux is

$$
F = |v|/V
$$

If we have two colliding beams, the probability of finding either particle is $1/V$ but since the collision can occur anywhere in the box, the total flux is

$$
|v_1 - v_2|/V^2 \times V = |v_1 - v_2|/V$$
Let us look in more detail at $2 \rightarrow 2$ scattering in the center of mass frame. The 3-velocities are $v_1 = p_1/E_1$ and $v_2 = p_2/E_2 = -p_1/E_2$ and so

$$|v_1 - v_2| = |p_1| \left( \frac{1}{E_1} + \frac{1}{E_2} \right) = p_1 \frac{E_1 + E_2}{E_1 E_2} = p_1 E_T E_1 / E_2$$

which leads together with the two-body phase space of Eq. (6.56) to

$$d\sigma = \frac{1}{4E_1 E_2} |v_1 - v_2| \cdot |A_{fi}|^2 \cdot d\Omega_2 = \frac{1}{4E_1 E_2} \frac{E_1 E_2}{p_f E_T} |A_{fi}|^2 \frac{1}{16\pi^2} \frac{p_f d\Omega}{E_T} = \frac{1}{4p_i E_T} \frac{1}{16\pi^2} \frac{p_f |A_{fi}|^2}{E_T} d\Omega$$

from which we obtain

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2} \frac{1}{E^2} \frac{p_f |A_{fi}|^2}{p_i}$$ (6.60)

where $p_i$ and $p_f$ are the magnitudes of the three-momenta of the incoming and outgoing particles.

### 6.10 Green’s Functions

Often times, we are not interested in questions that can be answered by calculating scattering cross-sections. We might want to figure out the non-Gaussianity of density perturbations arising in the CMB from novel models of inflation or calculate the optical properties of strange metals. These questions are best addressed with the help of correlation functions, which we will now learn how to calculate with Feynman diagrams.

#### 6.10.1 The true vacuum

So far we have approximated the ground state with the ground state of the free theory $H_0|0\rangle = 0$. We will now define the true vacuum of the interacting theory $|\Omega\rangle$. The energy is chosen such that for $H = H_0 + H_{int}$

$$H|\Omega\rangle = 0$$ (6.61)

and we normalize the vacuum to

$$\langle \Omega | \Omega \rangle = 1$$

We define correlation functions as

$$G^{(n)}(x_1, \ldots, x_n) \equiv \langle \Omega | T\phi_H(x_1) \cdots \phi_H(x_n) | \Omega \rangle$$ (6.62)

where $\phi_H$ is in the Heisenberg picture of the full theory. We also call $G^{(n)}$ the Green’s functions of the full theory. How can we compute these objects using Feynman Diagrams?
\textbf{Claim:} The $n$-point correlation function is given as
\begin{equation}
G^{(n)}(x_1, \ldots, x_n) = \langle \Omega | T \phi_H(x_1) \ldots \phi_H(x_n) | \Omega \rangle
\end{equation}
\begin{equation}
= \frac{\langle 0 | T \phi_I(x_1) \ldots \phi_I(x_n) S | 0 \rangle}{\langle 0 | S | 0 \rangle}
\end{equation}
where we the interaction picture operators on the right-hand side are applied on the vacuum of the free theory $|0\rangle$ and the usual time evolution operator
\begin{equation}
S = T \exp \left( -i \int dt \mathcal{H}_I(x) \right) = \lim_{t_+ \to +\infty \atop t_- \to -\infty} U(t_+, t_-)
\end{equation}
\textbf{Proof:} The time evolution operator in the interaction picture
\begin{equation}
U(t, t_0) = T \exp \left( -i \int_{t_0}^{t} dt' \int d^3 x \mathcal{H}_I(x, t') \right)
\end{equation}
has the following properties
\begin{equation}
U^\dagger(t, t_0) = U(t_0, t), \quad U(t_0, t_1)U(t_1, t_2) = U(t_0, t_2), \quad U(t_0, t_0) = 1
\end{equation}
The numerator of the RHS of Eq. (6.64) can be written with
\begin{equation}
\phi_H(x) = U^\dagger(t, t_0) \phi_I(x, t) U(t, t_0) \quad (\text{see Eq. (6.12))}
\end{equation}
as
\begin{equation}
\langle 0 | T \phi_I(x_1) \ldots \phi_I(x_n) U(t_+, t_-) | 0 \rangle
= \langle 0 | T U(t_0, t_1) \phi_I(x_1) U(t_1, t_0) U(t_0, t_2) \phi_H(x_2) U(t_2, t_0) \ldots U(t_0, t_n) \phi_H(x_n) U(t_n, t_0) U(t_+, t_-) | 0 \rangle
\end{equation}
Since we are allowed to commute operators under time ordering $T$, using $U(t_0, t_1)U(t_1, t_0) = 1$, we get
\begin{equation}
\langle 0 | T U(t_+, t_0) \phi_H(x_1) \phi_H(x_2) \ldots \phi_H(x_n) U(t_0, t_-) | 0 \rangle
\end{equation}
Since $t_\pm \to \pm \infty$ the ordering of the remaining $U$ is independent of the time-ordering $T$ and they will always appear to the left and right of the string of operators. We therefore need to figure out the meaning of the two boundary $U$ operators, e.g.
\begin{equation}
\lim_{t_- \to -\infty} U(t_0, t_-) | 0 \rangle
\end{equation}
We consider an arbitrary state $|\Psi\rangle$,
\begin{equation}
\langle \Psi | U(t_0, t_-) | 0 \rangle = \langle \Psi | U_H(t_0, t_-) | 0 \rangle
\end{equation}
Where we have used that $H_0 |0\rangle = 0$ and $U_H$ is the time evolution operator of the full theory. Let us insert a complete set of states, which we take as energy eigenstates of the full Hamiltonian $H = H_0 + H_{\text{int}}$,
\begin{equation}
\langle \Psi | U_H(t, -\infty) | 0 \rangle
= \lim_{t_- \to -\infty} \frac{\langle \Psi | U_H(t_0, t_-) | \Omega \rangle}{\langle \Omega | \Omega \rangle} + \sum_{n \neq 0} \frac{\langle n | \langle n | U_H(t_0, t_-) | 0 \rangle}{\langle n | n \rangle} | 0 \rangle
= \lim_{t_- \to -\infty} \langle \Psi | U_H(t_0, t_-) | \Omega \rangle | \Omega \rangle + \lim_{t_- \to -\infty} \sum_{n \neq 0} e^{iE_n(t_- - t)} \langle \Psi | n \rangle | n \rangle | 0 \rangle
= \langle \Psi | \Omega \rangle | \Omega \rangle + \lim_{t_- \to -\infty} \sum_{n \neq 0} e^{iE_n(t_- - t)} \langle \Psi | n \rangle | n \rangle | 0 \rangle
\end{equation}
Recall from Eq. (6.13), that in the interaction picture
\begin{equation}
U(t, t_0) = e^{iH_0(t-t_0)} e^{-iH(t-t_0)}
\end{equation}
and the full time-evolution is
\begin{equation}
U_H = e^{-iH(t-t_0)}
\end{equation}
where we have used Eq. (6.61) in the first term. The second term can also be simplified. In fact, it vanishes!

The sum \( \sum_n \) is actually an integral \( \int d^3 p \ldots \) since the momentum states form a continuum. We can now use the Riemann-Lebesgue lemma, which states that for any well-behaved function \( f(x) \),

\[
\lim_{\omega \to \infty} \int \alpha dx f(x) e^{i\omega x} = 0
\]

We can alternatively derive this by shifting the integration slightly into the imaginary \( t \to -\infty(1+i\varepsilon) \), which exponentially suppresses all energy eigenstates except for the ground-state.\(^{29}\) The above result therefore gives,

\[
\langle \Psi | U_H(t, -\infty) | 0 \rangle = \langle \Psi | \Omega \rangle \langle \Omega | 0 \rangle \quad (6.65)
\]

In particular, we see that the vacuum of the free theory and the vacuum of the full theory are proportional to each other

\[
| 0 \rangle = (\Omega | 0 \rangle \cdot | \Omega \rangle
\]

which justifies the somewhat hand-waving argument in Sec. 6.3.1 to use free states, or creation operators acting on the free vacuum \( | 0 \rangle \), in scattering processes.

If we insert the result of Eq. (6.65) in the RHS of Eq. (6.64), we find

\[
\frac{\langle 0 | T \phi_I(x_1) \ldots \phi_I(x_n) S | 0 \rangle}{\langle 0 | S | 0 \rangle} = \frac{\langle 0 | \Omega \rangle \langle \Omega | T \phi_H(x_1) \phi_H(x_2) \ldots \phi_H(x_n) | \Omega \rangle \langle \Omega | 0 \rangle}{\langle 0 | \Omega \rangle \langle \Omega | \Omega \rangle \langle \Omega | 0 \rangle} = \langle \Omega | T \phi_H(x_1) \phi_H(x_2) \ldots \phi_H(x_n) | \Omega \rangle
\]

where we have used \( \langle \Omega | \Omega \rangle = 1 \). This completes the proof of this very important result which we can write in slightly more explicit form

\[
\langle \Omega | T \phi_H(x_1) \ldots \phi_H(x_n) | \Omega \rangle = \frac{\langle 0 | T \phi_I(x_1) \ldots \phi_I(x_n) e^{-i \int dx H_I(x)} | 0 \rangle}{\langle 0 | e^{-i \int dx H_I(x)} | 0 \rangle} \quad (6.66)
\]

This expression for the \( n \)-point Green’s function \( G^{(n)}(x_1, \ldots, x_n) \) is ideally suited to doing perturbative calculations. We need only keep as many terms as needed in the Taylor series expansions of the exponentials.

### 6.10.2 Connected diagrams and vacuum bubbles

We will now sketch some further results before we move on to greener pastures\(^{30}\). With Dyson’s formula and Wick’s theorem, we know how to calculate the expressions appearing in Eq. (6.66). Feynman diagrams greatly simplify these calculations. But what is the meaning of the factor \( (\langle 0 | S | 0 \rangle)^{-1} \)? What are we dividing the

\[\text{Figure 6.14: The Riemann-Lebesgue lemma. The function } f(x) \text{ is multiplied by a rapidly oscillating function. The product integrates to zero in the limit of infinite oscillation frequency. Proof:}

\[
\int f(x) e^{-i \omega x} dx = \int f(x) e^{-i \omega x} dx \leq \frac{1}{\omega} \int |f'(x)| dx \to 0 \text{ as } \omega \to \pm \infty.
\]

Since the result is true for any state \( \langle \Psi | \), we have

\[
U_H(t, -\infty) | 0 \rangle = | \Omega \rangle \langle \Omega | 0 \rangle
\]

\[\text{With}

\[
\langle 0 | S | 0 \rangle = \lim_{t \to \mp \infty} \langle 0 | U(t_+, t_) | 0 \rangle
\]

\[= \lim_{t \to \mp \infty} \langle 0 | U(t_+, t) U(t, t_) | 0 \rangle
\]

\[= \lim_{t \to \mp \infty} \langle 0 | U(t_+, t) U(t, t-) | 0 \rangle
\]

\[= \langle 0 | \Omega \rangle \langle \Omega | \Omega \rangle \langle \Omega | 0 \rangle
\]

\[\text{Spin } \frac{1}{2} \text{ representations!}
\]
expression by? We will now sketch that the denominator removes all vacuum bubbles and that we only need to consider connected Feynman graphs, in words:

\[
\langle \Omega | T \phi_H(x_1) \ldots \phi_H(x_n) | \Omega \rangle = \sum \text{connected Feynman graphs}
\]  

(6.67)

where a connected diagram is defined by the requirement that every line is joined to an external leg.

Let us explain this in \( \phi^4 \) theory. The diagramatic expansion for \( \langle 0 | S | 0 \rangle \) is

\[
\langle 0 | S | 0 \rangle = 1 + \bigcirc + \left( \bigcirc + \bigcirc + \bigcirc \bigcirc \bigcirc \right) + \ldots
\]

None of these diagrams connect to external particles and we therefore call them vacuum bubbles. We will not show this here\(^{31}\), but the pre-factors of the diagrams consisting of combinatoric and symmetry factors conspire such that the whole series can be exponentiated

\[
\langle 0 | S | 0 \rangle = \exp \left( \bigcirc + \bigcirc + \bigcirc + \ldots \right)
\]

and we find that the amplitude for the vacuum of the free theory to evolve into itself (the numerator of the RHS of Eq. (6.66))

\[
\langle 0 | S | 0 \rangle = \exp(\text{all distinct vacuum bubbles})
\]

The same combinatoric simplification takes place for the numerator evaluating of Eq. (6.66) generic correlation functions. One can show

\[
\langle 0 | T \phi_I(x_1) \ldots \phi_I(x_n) | S | 0 \rangle = \langle 0 | S | 0 \rangle \cdot \sum \text{connected diagrams}
\]

where again connected means no vacuum bubbles, or that all the lines are joined to an external leg. We can now conclude that the division by \( \langle 0 | S | 0 \rangle \) has a very cute interpretation in terms of Feynman diagrams. We only need to calculate the connected Feynman graphs and we do not have to care for the vacuum bubbles. We therefore obtain as promised Eq. (6.67)

\[
\langle \Omega | T \phi_H(x_1) \ldots \phi_H(x_n) | \Omega \rangle = \sum \text{connected Feynman graphs}
\]

Here we show another example for \( \mathcal{H}_I(x) = \frac{g^2}{2} \phi^3(x) \) theory for a two-point function

\[
\langle 0 | T \phi_I(x_1) \phi_I(x_2) e^{-i \int d^4x \mathcal{H}_I(x)} | 0 \rangle = D_{12} - g^2 \int \left( \frac{1}{4} D_{1x} D_{2x} D_{y_1} D_{y_2} + \frac{1}{8} D_{12} D_{xx} D_{yy} + \frac{1}{12} D_{12} D_{xy}^3 + \ldots \right)
\]

See e.g. Schwartz 7.2.5 or Peskin 4.4. Compare also to the exponentiation of the cluster expansion in statistical physics, see script 3.4.1 of SS17 of the TH4A lecture by yours truly.
\[
\frac{D_{12}}{D_{1}, D_{2}, D_{\psi}, D_{\phi}} + \frac{D_{12} + D_{12}}{D_{12} + D_{12} + \ldots} = \left( \frac{D_{12}}{D_{12} + D_{12} + \ldots} \right) \times (1 + \frac{D_{12}}{D_{12} + D_{12} + \ldots} + \ldots)
\]

where integration over repeated indices is implied and \( D_F(x_1, x_2) = D_{12} \). The sum over all graphs (disconnected and connected) in the numerator is then the sum over all graphs with no bubbles multiplying the sum over the bubbles. In the calculation of \( \langle \Omega | T \phi_{H}(x_1) \ldots \phi_{H}(x_n) | \Omega \rangle \), we therefore can ignore all bubble diagrams, they are automatically included by using the real vacuum \( | \Omega \rangle \) instead of the free one \( | 0 \rangle \).

### 6.10.3 Outlook to S-matrices from Green’s functions

How can we related the Green’s functions back to S-matrix elements relevant for scattering processes? The Feynman rules for the Green’s functions include the propagators for the external legs, which we have to amputate. Also, the 4-momenta assignment to the external legs in \( \tilde{G}^{(n)} \) is arbitrary and generally not on-shell. In order to achieve both we can relate the Green’s function in momentum space \( \tilde{G}^{(n)}(p_1, \ldots, p_n) \) to S-matrix elements

\[
\langle q_1, \ldots, q_n | S - 1 | p_1, \ldots, p_{n'} \rangle = (-i)^{n+n'} \prod_{i=1}^{n}(q_i^2 - m^2 + i\varepsilon) \prod_{j=1}^{n'}(p_j^2 - m^2 + i\varepsilon) \tilde{G}^{(n+n')}(-q_1, \ldots, -q_n, p_1, \ldots, p_{n'})
\]

where we take the limit \( p^0 \rightarrow E_p \) in all of the external legs. These prefactors therefore vanish and we only get a non-zero answer for diagrams contributing to \( G^n \) which have the same propagators for each external leg.

The important step that we have taken here, is that this provides a systematic approach to dealing with true particle states in an interacting theory. This formula still applies once we take into account virtual particles in asymptotic states. This is the formally correct way to consider scattering and is known as the LSZ reduction formula. It will be worked out and shown in more detail in your quantum field theory lecture.
Figure 6.15: Self-energy-like diagrams representing 32 gauge-invariant subsets contributing to the muon $g - 2$ at the tenth order, which leads to prediction for the anomalous magnetic dipole moment $a = \frac{2}{\pi} g$ as

$$a_{\mu}(\text{SM}) = 116,591,840 \times 10^{-11}$$

and compare to the measurement

$$a_{\mu}(\text{exp}) - a_{\mu}(\text{SM}) = 249 \times 10^{-11}$$

Lorentz Representations

In non-relativistic quantum mechanics we learn that the spin \( \pm \frac{1}{2} \) states of the electron along a specific direction, are conveniently described by a complex doublet of states

\[
|\psi\rangle = \begin{pmatrix} \psi_\uparrow \\ \psi_\downarrow \end{pmatrix}
\]

(7.1)

The dynamics of the doublet \( \psi = \langle x | \psi \rangle \) is governed by the Schrödinger-Pauli equation

\[
i \partial_t \psi = \left[ \left( \frac{1}{2m} (i \nabla - e A)^2 - e A_0 \right) \mathbb{1}_{2 \times 2} + \mu_B B \cdot \sigma \right] \psi
\]

(7.2)

with the vector potential \( A_{\mu} \), \( B = \nabla \times A \), and the Bohr magneton

\[
\mu_B = \frac{e}{2m_e}
\]

describing the strength of the coupling between the \( B \)-field and the electron’s magnetic moment.\(^1\)

The 3 \( \sigma \) = \( (\sigma_1, \sigma_2, \sigma_3) \) are the Pauli-matrices which we have combined into a vector:

\[
\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
\]

(7.3)

Now it would seem that we have a problem with rotational invariance, since the \( \sigma \)-vector will not change under rotations contrary to \( \psi \) and \( B \). How do we make sure that the equation is rotationally invariant? It works because

\[
[\sigma_i, \sigma_j] = 2i \varepsilon_{ijk} \sigma_k
\]

are the exact same algebraic relations that the generator of infinitesimal rotations satisfy, see Eq. (7.8) and the discussion below. We anticipate the result and conclude that the changes induced by rotations cancel in

\[
B \cdot \sigma \psi
\]

or any other vector field, like \( \partial_i \), which transforms also as a 3-vector. So we could have written down a rotationally invariant equation for \( \psi \) with

\[
(1 \partial_t - \partial_i \sigma_i)\psi = 0
\]

(7.4)

\(^1\) Recall the Stern-Gerlach experiment.
It turns out, this equation is Lorentz-invariant, too. We combine
\[ \sigma^\mu = (1, \sigma_1, \sigma_2, \sigma_3) \] (7.5)
to obtain
\[ \sigma^\mu \partial_\mu \psi = 0 \] (7.6)
This is the **Dirac equation**! We could try to be bold and add a mass term with
\[ (\sigma^\mu \partial_\mu + m)\psi = 0 \] (7.7)
but this is not a Lorentz invariant equation. Please forget Eq. (7.7)
immediately (and remember only Eq. (7.6)).

In conclusion, we need to understand the intricate Lorentz props-
ties of spin-1/2 particles, discuss their representations. The Dirac
equation and its non-relativistic limit, the Schrödinger-Pauli equa-
tion, will follow.

7.1 The Lorentz algebra

We will now be a bit more rigorous and introduce groups, algebras
and representations. A **group** consists of a set of elements \( \{g_i\} \) and
a rule
\[ g_i \times g_j = g_k \]
The rule is required to be associative and have an identity element.
A **representation** is a particular embedding of the \( g_i \) into operators
that act in a vector space. For finite-dimensional representations this
means an embedding of the \( g_i \) into matrices.

Every group has a trivial representation \( r : g_i \rightarrow 1 \). We dis-
tinguish this from a **faithful representation**, where each group
element has its own matrix.

For the Lorentz-group, which we recall from Eq. (3.58) leaves
\[ \eta_{\alpha\beta} = \eta_{\mu\nu} \Lambda^\mu_{\;\alpha} \Lambda^\nu_{\;\beta} \]
invariant, we have only discussed the fundamental representation
which acts on 4-vectors
\[ x^\mu \rightarrow \Lambda^\mu_{\;\nu} x^\nu \]
and if you go back to Sec. 3.3.1, you’ll find examples of this embed-
ding as the set of 4 \times 4 rotation and boost matrices. Our goal is now
to find all the representations.

Our goal is now to extract the properties of the group inde-
pendent of one particular representation. We will now focus on
continuously connected groups, which can be described by a set
of coordinates which are real numbers. The easiest approach is to
consider **infinitesimal transformations**: for any group \( G \), group
elements \( g \in G \) can be written as
\[ g = \exp(i\alpha^g_{\lambda_i}) \]

This is the equation for a **Weyl** 2-spinor, which is a slightly differ-
ent form of the equation which we usually associate with the Dirac equa-
tion, which describes the dynamics of a **Dirac** 4-spinor.

Think of e.g. matrices acting on vectors. Technically, the matrix
embedding is the representation, not the vectors, even though we are
not careful usually in making this distinction.

A non-continuous group, is e.g. the reflection group \( Z_2 \), which has the
elements 1, −1
where the $\alpha_i^g$ are just numbers and the $\lambda_i$ are group generators. Generators $\lambda_i$ form an algebra, because we can add and multiply them, whereas group elements can only be multiplied. They can be extracted from the group elements using infinitesimal transformations
\[
g = \exp(i\alpha_i^g \lambda_i) = 1 + i\alpha_i^g \lambda_i + \ldots
\]
or through
\[
\lambda_i = \frac{1}{i} \left. \frac{dg}{d\alpha_i^g} \right|_{\alpha_i^g = 0}
\]
**Lie groups** are a special class, with an infinite number of elements but with a finite number of generators. The generators of a Lie group form the Lie algebra. Lie groups are crucial for the SM as it contains the $SU(3) \times SU(2) \times U(1)$ special unitary groups. The Lorentz group is also a (non-compact) Lie group and is sometimes called $O(1,3)$, an orthogonal group that preserves vector products with a metric with the $(1,3)$ signature $\eta_{\mu\nu} = (1,-1,-1,-1)$.

Let us start with a reminder of the rotation group $O(3)$. The three generators $J_i$ of the rotation group satisfy
\[
[J_i, J_j] = i\varepsilon_{ijk} J_k
\] (7.8)
when we act on a space-time vector $x^\mu$ they are represented as
\[
J_1 = \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & -i \\
i & 0 & 0 & 0
\end{pmatrix}, \quad J_2 = \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & i \\
0 & 0 & 0 & 0 \\
0 & -i & 0 & 0
\end{pmatrix}, \quad J_3 = \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & -i & 0 \\
i & 0 & 0 & 0
\end{pmatrix}
\]
and so e.g. for the group element $G_1 = \exp(i\theta J_1)$
\[
G_1 = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & \cos \theta & \sin \theta \\
0 & 0 & -\sin \theta & \cos \theta
\end{pmatrix}
\] (7.9)
You can obtain $J_2$ and $J_3$ by straight-forward permutations. How about the Lorentz-boosts? We recall Eq. (3.17)
\[
\begin{pmatrix}
x' \\
t'
\end{pmatrix} = \begin{pmatrix}
\cosh \eta & \sinh \eta \\
\sinh \eta & \cosh \eta
\end{pmatrix} \begin{pmatrix}
x \\
t
\end{pmatrix}
\] (7.10)
which means infinitesimally
\[
x \rightarrow x + \eta t \\
t \rightarrow t + \eta x
\]
The generator of a Lorentzboost in the $x$ direction
\[
\begin{pmatrix}
t \\
x \\
y \\
z
\end{pmatrix} = \begin{pmatrix}
t \\
x \\
y \\
z
\end{pmatrix} + iK_1 \begin{pmatrix}
x \\
y \\
z
\end{pmatrix}
\]
Think of the rotation group in 3D, which contains infinitely many rotations matrices depending on the 3 continuous Euler rotation angles, but there are only 3 different infinitesimal rotations in the $x - y$, $x - z$, and $y - z$ plane.
is therefore given by the hermitian matrix

\[ iK_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \]

and analogously

\[ iK_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad iK_3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \]

We can now check the commutator between the generators of rotations \( J_i \) and the boost generators \( K_i \). We find

\[ [J_i, K_j] = i\varepsilon_{ijk}K_k \quad (7.11) \]

What does that mean? This just reflects that the boost generators \( K_i \) transform as a 3-vector under rotations, as expected since the boost direction is a 3-vector characterized by the relative velocity \( v^i \).

We will now do one of the most important calculations in history. Under a rotation we would get for the boost group elements \( U = \exp(i\eta^i K_i) \) which are \( 4 \times 4 \) matrices,

\[ U \rightarrow G^{-1} U G = (1 - iJ_i)(1 + i\eta^i K_i)(1 - iJ_i) + \ldots \]

Under a rotation we would get for the boost group elements \( U = \exp(i\eta^i K_i) \) which are \( 4 \times 4 \) matrices,

\[ U \rightarrow G^{-1} U G = (1 - iJ_i)(1 + i\eta^i K_i)(1 - iJ_i) + \ldots \]

and so

\[ K_i \rightarrow i[J_i, K_i] \]

Two boosts produce a rotation! Summarizing, we find the Lorentz-algebra \( SO(1,3) \)

\[ [J_i, J_j] = i\varepsilon_{ijk}J_k \quad (7.12) \]
\[ [J_i, K_j] = i\varepsilon_{ijk}K_k \quad (7.13) \]
\[ [K_i, K_j] = -i\varepsilon_{ijk}J_k \quad (7.14) \]

Please note the very important negative sign in the commutator of two boosts!

Now use a crucial observation: the algebra falls apart in two independent sets if complexify it using

\[ J_i^+ \equiv \frac{1}{2}(J_i + iK_i), \quad J_i^- \equiv \frac{1}{2}(J_i - iK_i) \]

which satisfies

\[ [J_i^+, J_j^+] = i\varepsilon_{ijk}J_k^+ \quad (7.15) \]
\[ [J_i^-, J_j^-] = i\varepsilon_{ijk}J_k^- \quad (7.16) \]
\[ [J_i^+, J_j^-] = 0 \quad (7.17) \]
This shows that the complexified Lie-algebra for the Lorentz-group has two commuting sub-algebras. The algebra generated by $J^\pm$ is the 3D rotation algebra or so(3) = sl(2, R) = su(2). So we have shown that

$$\text{so}(1, 3) = \text{su}(2) \oplus \text{su}(2)$$

which means that the Lorentz group $SO(1, 3)$ is locally isomorphic to two independent $SU(2)$ groups. As one can show, the fact that the complexifications of the corresponding Lie algebras coincide implies that the representation theory of both groups coincides.

### 7.1.1 Representations

We can now simply use our knowledge of $SU(2)$ representations to determine all representations of $SO(1, 3)$. We know that the representations are labeled by their spin

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$$

and each $SU(2)$ representation consists of $(2j + 1)$ objects

$$\psi_m \quad \text{with} \quad m = -j, -j + 1, \ldots, j - 1, j$$

which are rotated into each other by elements of $SU(2)$. We therefore know that the irreducible representations of $SO(1, 3)$ are labeled by $(j^+, j^-)$ with $j^\pm$ taking on the values of Eq. (7.18).

Each $SO(1, 3)$ representation contains $(2j^+ + 1)(2j^- + 1)$ objects $\Psi_{m^+, m^-}$ with

$$\Psi_{m^+, m^-} \quad \text{with} \quad m^\pm = -j^\pm, -j^\pm + 1, \ldots, j^\pm - 1, j^\pm$$

<table>
<thead>
<tr>
<th>$SU(2) \otimes SU(2)$</th>
<th>(0, 0)</th>
<th>$(\frac{1}{2}, 0)$</th>
<th>$(0, \frac{1}{2})$</th>
<th>$(\frac{1}{2}, \frac{1}{2})$</th>
<th>(1, 0)</th>
<th>(1, 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$SO(3)$</td>
<td>0</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
<td>0 $\oplus$ 1</td>
<td>1</td>
<td>0 $\oplus$ 1 $\oplus$ 2</td>
</tr>
</tbody>
</table>

We recognize the trivial, one-dimensional representation $(0, 0)$ as the Lorentz-scalar. Counting the size of the representation, we conclude that the 4-dimensional representation $(\frac{1}{2}, \frac{1}{2})$ is the Lorentz vector, which is the defining representation of the Lorentz group.

### 7.1.2 Spinor representations

How can we understand the representations $(\frac{1}{2}, 0)$? We write

$$\psi_L = \psi_\alpha, \quad \text{with} \quad \alpha = 1, 2$$

How about the notation $(\frac{1}{2}, 0)$? This says that $J_i^+ = \frac{1}{2}(J_i + iK_i)$ acting on $\psi_\alpha$ is represented by

$$\frac{1}{2}\sigma_i$$

Table 7.1: Decomposition of irreducible representations of the Lorentz $SU(2) \times SU(2)$ into irreducible representations of its $SO(3)$ subgroup describing spin.

The name "left-handed" and "right-handed" from the way the spin precesses as a massless fermion moves: we will see this in Sec. 7.3.4.
while \( J_i^- = \frac{1}{2} (J_i - iK_i) \) is represented by 0. This means that the representations of the generators (before complexification) are

\[
J_i = \frac{1}{2} \sigma_i \\
iK_i = \frac{1}{2} \sigma_i
\]

after solving the simple linear system. The equality here means "represented by".

We will similarly denote the \((\frac{1}{2}, 0)\) with the slightly exotic symbol

\[
\psi_R = \bar{\chi}^\dot{\alpha}, \quad \text{with} \quad \dot{\alpha} = 1, 2
\]

(7.20)

Do not be confused by the bar. For now it is just part of the typographic symbol.

The two-component spinors \( \psi_\alpha \) and \( \bar{\chi}^{\dot{\alpha}} \) are called \textbf{Weyl spinors} and furnish perfectly good representations of the Lorentz group.

What about the \textbf{Dirac spinor}? If you have already encountered it, you know that it has also 4 components. Why? The reason is \textbf{parity}. We know that under parity, see XXXXX, that

\[
x \rightarrow -x \\
t \rightarrow t
\]

and similarly for other 3-vectors like the 3-momentum or the boosts

\[
p \rightarrow -p \\
K \rightarrow -K
\]

Rotations are \textbf{pseudo-vectors} and they transform as

\[
J \rightarrow J
\]

This means for the complexified generators

\[
J_i^+ = \frac{1}{2} (J_i + iK_i) \rightarrow \frac{1}{2} (J_i - iK_i) = J_i^- 
\]

In other words \( J^+ \leftrightarrow J^- \) and so under parity we exchange

\[
(\frac{1}{2}, 0) \leftrightarrow (0, \frac{1}{2})
\]

In order to describe an electron in QED which conserves parity, we need to use both 2-dimensional representations, which we can describe mathematically

\[
(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})
\]

This \textbf{reducible} representation is thus given by stacking two Weyl spinors to form a Dirac spinor

\[
\Psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = \begin{pmatrix} \psi_\alpha \\ \bar{\chi}^{\dot{\alpha}} \end{pmatrix}
\]

We use dotted and undotted indices to make sure we don’t accidentally contract the \( \alpha \) of \((\frac{1}{2}, 0)\) with an \( \dot{\alpha} = 1, 2 \) of \((\frac{1}{2}, 0)\), which would not be Lorentz invariant. This is called van der Waerden notation. It will prove extremely useful in supersymmetry and in superstring theory.

Think of the angular momentum as \( L = r \times v \) or imagine a clock-wise rotation in a mirrored coordinate system.

QED and QCD are parity invariant, which means the equations of motions and the Lagrangian are invariant under \( P \) transformations. If had chosen to write the theory using just the \((\frac{1}{2}, 0)\) spinor, we would have explicitly violated \( P \). This is in fact, what happens in the electro-weak sector: \( SU(2)_L \) gauge bosons only couple to \((\frac{1}{2}, 0)\) spinors and not to the \((0, \frac{1}{2})\) representation.
The spinor $\Psi(p)$ is (like the scalar) a function of 4-momentum, or a function of $x^\mu$ if you Fourier-transform. We know its transformation properties under rotations

$$ J_{\text{Dirac}} = \begin{pmatrix} \frac{i}{2} \sigma & 0 \\ 0 & \frac{i}{2} \sigma \end{pmatrix} $$

where the equality means "represented by". Similarly for boosts

$$ iK_{\text{Dirac}} = \begin{pmatrix} \frac{i}{2} \sigma & 0 \\ 0 & -\frac{i}{2} \sigma \end{pmatrix} $$

Note, the very important minus sign, which distinguishes the two irreducible representations entering a Dirac spinor.

Parity forced us to use a 4-spinor but we know that a electron has only two physical degrees of freedom. We must project out two of four the components described by $\Psi(p)$. Let us go to the rest-frame

$$(p^\mu)_r \equiv (m, 0)$$

Anticipating the result, we will understand fully later, we write the projection operator as

$$ P = \frac{1}{2}(1 - \gamma^0) $$

For now $\gamma^0$ is just a $4 \times 4$ matrix. A projection applied twice should be the same as just applying it once

$$ P^2 = P \quad \Rightarrow \quad (\gamma^0)^2 = 1 $$

and the eigen-values of $\gamma^0$ are therefore $\pm 1$. Since we know that under parity

$$ \psi_L \leftrightarrow \psi_R $$

we cannot simply use the projection operator to set one of the fields (e.g. $\bar{\chi}^\alpha$) to 0. Parity means that the two irreducibles should be on the same footing. We choose the following representation of $\gamma^0$

$$ \gamma^0 = \begin{pmatrix} 0 & 1_{2 \times 2} \\ 1_{2 \times 2} & 0 \end{pmatrix} $$

or explicitly

$$ \gamma^0 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} $$

So in the rest-frame, we can write the projection to two degrees of freedom as

$$ P \Psi(p_r) = 0 $$

which is the same as

$$ \psi_L - \psi_R = 0 \quad (7.21) $$
Since this is an equation equating the two-spinors, it removes two of the four degrees of freedom contained in the Dirac spinor.

Eq. (7.21) does not look Lorentz invariant since it equates states in two different representations. Note however, that we required a specific frame (the rest frame) and therefore should not respect the resulting equation to be covariant. Once we boost to a generic frame, we will find that the projection equation is covariant.

7.1.3 The Dirac equation

We have "derived" the Dirac equation above. Why? We know now how to obtain the equation which is satisfied by \( \Psi(p) \) of a general \( p^\mu \) by Lorentz-transforming the spinor! We boost

\[ \Psi(p) = e^{-i\eta \cdot K} \Psi(p_{r}) \]

which implies for the projection equation for a general \( p^\mu \)

\[ \left( e^{-i\eta \cdot K} \gamma^0 e^{i\eta \cdot K} - 1 \right) \Psi(p) \]

we define

\[ \frac{\gamma^\mu p_\mu}{m} \equiv e^{-i\eta \cdot K} \gamma^0 e^{i\eta \cdot K} \]

to obtain the Dirac-equation

\[ (\gamma^\mu p_\mu - m)\Psi(p) = 0 \] (7.22)

which we can write in coordinate space as

\[ (i\gamma^\mu \partial_\mu - m)\Psi(x) = 0 \] (7.23)

which is a first-order differential equation for the Dirac 4-spinor.

We can expand on the last steps by noting that

\[ e^{-i\eta \cdot K} \gamma^0 e^{i\eta \cdot K} = \begin{pmatrix} 0 & \exp(-\eta \cdot \sigma) \\ \exp(\eta \cdot \sigma) & 0 \end{pmatrix} \]

and with the unit vector \( \eta/|\eta| \) we get

\[ \exp(\eta \cdot \sigma) = \cosh |\eta| + \frac{\eta}{|\eta|} \cdot \sigma \sinh |\eta| \]

If we identify the relation between boost and the 3-momentum and the relevant boost as

\[ p = m \frac{\eta}{|\eta|} \sinh |\eta| \] (7.24)

We can then easily see that

\[ \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \]

The Dirac matrices satisfy the anti-commutator relations

\[ \{ A, B \} = AB + BA \]
\[
\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu} 1_{4 \times 4}
\] (7.25)

The \(\gamma\)-matrices generate the **Dirac algebra**, which is a special case of a so-called Clifford algebra. This particular basis is the the **Weyl-representation**.

We additionally define

\[
\sigma^\mu \equiv (1, \sigma), \quad \bar{\sigma}^\mu \equiv (1, -\sigma)
\] (7.26)

and we can write the Dirac matrices in the **Weyl or chiral** basis as

\[
\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix}
\] (7.27)

We have found a non-trivial representation of the Lorentz group. Note however that the representation is **non-unitary**! The complexified algebra is unitary but it will come with imaginary coefficients.

You can see that most easily by looking at an example of a boost in the \(z\)-direction for \((\frac{1}{2}, 0)\)-representation which generated by \(K_3\)

\[
iK_3 = \frac{\sigma_3}{2} \quad \Rightarrow \quad K_3 = \frac{1}{2} \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}
\]

which means that the group \(g_3(\eta_3) \in SO(1, 3)\) element is

\[
g_3(\eta_3) = \exp(i\eta_3 K_3) = \begin{pmatrix} e^{\frac{\eta_3}{2}} & 0 \\ 0 & e^{-\frac{\eta_3}{2}} \end{pmatrix}
\]

This is **not** a unitary matrix, since \(g_3^\dagger g_3 \neq 1\). This is expected since the Lorentz group is **not compact**, and therefore there are no finite dimensional unitary representations of the Lorentz group. We will have to make use of projective representations and work with a so-called little group, but this is for your next QFT course.

We can find a Lorentz-covariant form of the Lorentz-algebra in Eq. (7.14) by generalizing the rotation group generators and writing them as differential operators

\[
J = x \times p = x \times (-i\nabla)
\]

which we can also write as an anti-symmetric tensor with \(J^3 = J^{12}\) and so on:

\[
J^3 = -i(x^i\nabla^j - x^j\nabla^i)
\]

The generalization to 4 dimensions is naturally

\[
J^{\mu\nu} = i(x^\mu \partial^\nu - x^\nu \partial^\mu)
\] (7.28)

which is an anti-symmetric \(4 \times 4\) matrix, which contains \(N(N-1)/2 = 3 \cdot 4/2 = 6\) elements. We can identify them with the previously defined \(J_i\) and \(K_i\) by

\[
J^i \rightarrow J^{ij} \quad \text{and} \quad K^i \rightarrow J^{0i}
\]

---

3 The rotational subgroup is compact since it is parametrized by angles \(\theta_i\) which lie on a circle \(S^1\) with 0 and \(2\pi\) identified (there is more to this but since the topology of \(SO(3)\) is a ball with antipodal surface points identified). **Boosts** however are parametrized by \(\eta = v/c\) with \(0 \leq \eta < 1\) where \(\eta = 0\) and \(\eta = 1\) are not the same and the group is therefore not compact.

4 See e.g. S. Weinberg, QFT1, chapter 2.
again with \( J^3 = J^{12} \) and so on. A general Lorentz transformation is parametrized by the anti-symmetric tensor

\[
\omega_{\mu\nu} = -\omega_{\nu\mu}
\]

and

\[
\Lambda = e^{-\frac{i}{2}\omega_{\mu\nu}J^{\mu\nu}}
\]

The commutation rules of the generators can be determined simply from the commutators of the differential operators\(^5\)

\[
[J^{\mu\nu}, J^{\rho\sigma}] = i(\eta^{\nu\rho} J^{\mu\sigma} - \eta^{\mu\rho} J^{\nu\sigma} - \eta^{\nu\sigma} J^{\mu\rho} + \eta^{\mu\sigma} J^{\nu\rho}). \tag{7.29}
\]

You can see that the fundamental representation for \( J^{\mu\nu} \) acting on 4-vectors is given by

\[
(J^{\mu\nu})_{\alpha\beta} = i(\delta^{\mu\alpha}\delta^{\nu\beta} - \delta^{\mu\beta}\delta^{\nu\alpha}) \tag{7.30}
\]

Compare with the definition of \( K_i \) and \( J_i \) above to confirm. We can understand it better by looking at an infinitesimal transformation of \( x^\mu \)

\[
x^\alpha \to \left( \delta^\alpha_\beta - \frac{i}{2} \omega_{\mu\nu}(J^{\mu\nu})^\alpha_\beta \right) x^\beta
\]

which becomes e.g. in the case of \( \omega_{23} = -\omega_{32} = \theta \)

\[
x^\alpha \to \begin{pmatrix} 1 & 1 & -\theta \\ 1 & 1 & \theta \\ \theta & -\theta & 1 \end{pmatrix} x^\nu
\]

or in exponentiated form \( e^{-i\omega_{23} J^{23}} \), we reproduce Eq. (7.9)

\[
G_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos \theta & \sin \theta \\ 0 & 0 & -\sin \theta & \cos \theta \end{pmatrix} \tag{7.31}
\]

and you can also verify that \( \omega_{10} = -\omega_{01} = \eta \) gives \( K_1 \).

The Lorentz-algebra for Dirac spinors can be represented by

\[
S^{\mu\nu} = \frac{i}{4} [\gamma^\mu, \gamma^\nu] \tag{7.32}
\]

which can be written with Eq. (7.25) as

\[
S^{\mu\nu} = \frac{i}{4} [\gamma^\mu, \gamma^\nu] = \begin{cases} 0 & \mu = \nu \\ \frac{1}{2} \gamma^\mu \gamma^\nu & \mu \neq \nu \end{cases} \tag{7.33}
\]

\[
= \frac{i}{2} \gamma^\mu \gamma^\nu - \frac{i}{2} \eta^{\mu\nu} \tag{7.34}
\]

you can check that Eq. (7.32) is a representation of the Lorentz-algebra satisfying Eq. (7.29)

\[
[S^{\mu\nu}, S^{\rho\sigma}] = i(\eta^{\nu\rho} S^{\mu\sigma} - \eta^{\mu\rho} S^{\nu\sigma} - \eta^{\nu\sigma} S^{\mu\rho} + \eta^{\mu\sigma} S^{\nu\rho}). \tag{7.35}
\]
by using Eq. (7.25) repeatedly. So for a given Lorentz-transformation \( \Lambda^\mu_\nu \) defined by the parameters \( \omega_{\mu\nu} \), we have the Dirac-spinor representation

\[
S[\Lambda] = \exp \left( -\frac{i}{2} \omega_{\mu\nu} S^{\mu\nu} \right) \tag{7.36}
\]

with the transformation

\[
\Psi(x) \rightarrow S[\Lambda] \Psi(\Lambda^{-1}x)
\]

We can again see, that the Lorentz-transformations are not unitary, since the exponent is not hermitian\(^6\) because

\[
(S^{\mu\nu})^\dagger = -\frac{i}{4} [(\gamma^\nu)^\dagger, (\gamma^\mu)^\dagger]
\]

which would be hermitian if all the \( \gamma^\mu \) are hermitian \( (\gamma^\mu)^\dagger = \gamma^\mu \), or all are anti-hermitian \( (\gamma^\mu)^\dagger = -\gamma^\mu \). However this never happens because of Eq. (7.25)

\[
(\gamma^0)^2 = 1 \quad \Rightarrow \quad \text{real eigenvalues}
\]

\[
(\gamma^i)^2 = -1 \quad \Rightarrow \quad \text{imaginary eigenvalues}
\]

So we could pick \( \gamma^0 \) to be hermitian, but we can only chose \( \gamma^i \) to be anti-hermitian. Indeed, you can check that the Dirac matrices in the Weyl or chiral basis of Eq. (7.27) exactly satisfy this

\[
(\gamma^0)^\dagger = \gamma^0, \quad (\gamma^i)^\dagger = -\gamma^i
\]

We see that there is no way to pick the \( \gamma^\mu \) such that the exponent is hermitian and the representation is therefore non-unitary.

### 7.1.4 Lorentz-invariant actions

We have now a new field that we can work with, the Dirac spinor \( \Psi \). We now want to construct a Lorentz invariant action, which will lead to a Lorentz invariant equation of motion.

Let us try something naive, which will not work. We define the adjoint spinor as

\[
\Psi^\dagger(x) = (\Psi^\ast)^T(x)
\]

Can we build a Lorentz scalar by contracting \( \Psi^\dagger \Psi = \sum_{i=0}^3 (\Psi^\ast)_i \Psi_i \)? How does this behave under Lorentz transformations

\[
\Psi(x) \rightarrow S[\Lambda] \Psi(\Lambda^{-1}x)
\]

\[
\Psi^\dagger(x) \rightarrow \Psi^\dagger(\Lambda^{-1}x) S[\Lambda]^\dagger
\]

from which we obtain

\[
\Psi^\dagger(x) \Psi(x) \rightarrow \Psi^\dagger(\Lambda^{-1}x) S[\Lambda]^\dagger S[\Lambda] \Psi(\Lambda^{-1}x)
\]

\(^6\) For general unitary matrix

\[
U = \exp(iA)
\]

in order to satisfy

\[
U^\dagger U = 1
\]

we see

\[
U^\dagger U = \exp(-iA^\dagger) \exp(iA)
\]

\[
= 1 - i(A^\dagger - A) + \ldots
\]

we need

\[
A = A^\dagger
\]
This would be invariant if only $S[\Lambda]$ was unitary but we in general have $S[\Lambda]^\dagger S[\Lambda] \neq 1$. This means $\Psi^\dagger \Psi$ is not a Lorentz-invariant scalar, and certainly cannot be used in an action.

But we can use our insight in why it fails to find a solution. We pick a representation of the Clifford algebra which like the Weyl basis satisfies $(\gamma^0)^\dagger = \gamma^0$ and $(\gamma^\mu)^\dagger = -\gamma^\mu$, then we have

$$\gamma^0 \gamma^\mu \gamma^0 = (\gamma^\mu)^\dagger$$

which means that

$$(S^{\mu\nu})^\dagger = \frac{i}{4} [ (\gamma^\mu)^\dagger, (\gamma^\nu)^\dagger ] = \gamma^0 S^{\mu\nu} \gamma^0$$

such that

$$S[\Lambda]^\dagger = \exp \left( \frac{i}{2} \omega_{\mu\nu} (S^{\mu\nu})^\dagger \right) = \gamma^0 (S[\Lambda]^{-1}) \gamma^0 \quad (7.37)$$

This allows us to define the Dirac adjoint

$$\overline{\Psi}(x) = \Psi^\dagger(x) \gamma^0 \quad (7.38)$$

We now see that $\overline{\Psi} \Psi$ is a Lorentz scalar since with Eq. (7.37)

$$\overline{\Psi}(x) \Psi(x) = \Psi^\dagger(x) \gamma^0 \Psi(x)$$

$$\rightarrow \Psi^\dagger(\Lambda^{-1} x) S[\Lambda]^\dagger \gamma^0 S[\Lambda] \Psi(\Lambda^{-1} x)$$

$$= \Psi^\dagger(\Lambda^{-1} x) \gamma^0 \Psi(\Lambda^{-1} x)$$

$$= \overline{\Psi}(\Lambda^{-1} x) \Psi(\Lambda^{-1} x)$$

which is indeed the transformation law for a Lorentz-scalar field, see e.g. Eq. (4.35).

Let us show another important bi-linear: $\overline{\Psi} \gamma^\mu \Psi$ is a Lorentz-vector. We claim, it therefore transforms as

$$\overline{\Psi}(x) \gamma^\mu \Psi(x) \rightarrow \Lambda^\mu_\nu \overline{\Psi}(\Lambda^{-1} x) \gamma^\nu \Psi(\Lambda^{-1} x)$$

and we treat the $\mu = 0, 1, 2, 3$ label on the $\gamma$ matrices as a true vector index. We can use this bi-linear to form Lorentz-scalars by contracting it with other Lorentz-vectors. We suppress the $x$-argument for notational simplicity.

$$\overline{\Psi} \gamma^\mu \Psi \rightarrow \overline{\Psi} S[\Lambda]^{-1} \gamma^\mu S[\Lambda] \Psi$$

To show that $\overline{\Psi} \gamma^\mu \Psi$ transforms as a vector, we need

$$S[\Lambda]^{-1} \gamma^\mu S[\Lambda] = \Lambda^\mu_\nu \gamma^\nu \quad (7.39)$$

which we show by working infinitesimally

$$\Lambda = 1 - \frac{i}{2} \omega_{\rho\sigma} J^{\rho\sigma} + \ldots$$

$$S[\Lambda] = 1 - \frac{i}{2} \omega_{\rho\sigma} S^{\rho\sigma} + \ldots$$
From Eq. (7.39) we find that we have to show
\[ [S^\rho_\sigma, \gamma^\mu] = -(J^{\rho\sigma})^\mu_\nu \gamma^\nu \] (7.40)
which holds if we plug in the definition of \( J^{\rho\sigma} \). We use
\[ \gamma^\mu \gamma^\nu = -\gamma^\nu \gamma^\mu + \{\gamma^\mu, \gamma^\nu\} \]
\[ = -\gamma^\nu \gamma^\mu + 2\eta^{\mu\nu} \]
to obtain for \( \rho \neq \sigma \) with Eq. (7.33) for the LHS
\[ [S^\rho_\sigma, \gamma^\mu] = \frac{i}{2} [\gamma^\rho \gamma^\sigma, \gamma^\mu] \]
\[ = \frac{i}{2} (\gamma^\rho \gamma^\sigma \gamma^\mu - \gamma^\mu \gamma^\rho \gamma^\sigma) \]
\[ = \frac{i}{2} (\gamma^\rho \{\gamma^\sigma, \gamma^\mu\} - \gamma^\rho \gamma^\mu \gamma^\sigma - \{\gamma^\mu, \gamma^\rho\} \gamma^\sigma + \gamma^\rho \gamma^\mu \gamma^\sigma) \]
\[ = i(\gamma^\rho \eta^{\rho\mu} - \gamma^\sigma \eta^{\rho\mu}) \] (7.41)
For the RHS we plug in Eq. (7.30) with \( (J^{\rho\sigma})^\mu_\nu = i(\eta^{\rho\mu} \delta^\sigma_\nu - \eta^{\rho\nu} \delta^\sigma_\mu) \)
to finally obtain
\[ -(J^{\rho\sigma})^\mu_\nu \gamma^\nu = -i(\eta^{\rho\mu} \gamma^\sigma - \eta^{\rho\sigma} \gamma^\mu) \]
which equals Eq. (7.41) and therefore proves Eq. (7.40).

Similarly, \( \overline{\Psi} \gamma^\mu \gamma^\nu \Psi \) transforms as a Lorentz-tensor.

7.1.5 The Dirac Lagrangian

We can now use the Lorentz-covariant bilinears \( \overline{\Psi} \Psi \), \( \overline{\Psi} \gamma^\mu \Psi \), and \( \overline{\Psi} \gamma^\mu \gamma^\nu \Psi \) to build Lorentz-invariant action. We choose
\[ S = \int d^4x \overline{\Psi}(x)(i\gamma^\mu \partial_\mu - m)\Psi(x) \]
(7.42)

This is the **Dirac action**. If we vary with respect to \( \overline{\Psi} \), we get
\[ (i\gamma^\mu \partial_\mu - m)\Psi(x) = 0 \]
which is the beautiful **Dirac equation**. It is first order in derivatives but miraculously Lorentz invariant. If we tried a similar thing with a scalar field it would look like \( \nu^\mu \partial_\mu \phi = \ldots \) which necessarily involves a privileged vector \( \nu^\mu \) in space-time and would not be Lorentz invariant.

The Dirac equation mixes the various components of \( \Psi \) through the \( \gamma^\mu \) matrices. Note, each individual component itself solves the Klein-Gordon equation. We write
\[ (i\gamma^\nu \partial_\nu + m)(i\gamma^\mu \partial_\mu - m)\Psi(x) = (-\gamma^\nu \gamma^\mu \partial_\nu \partial_\mu + m^2)\Psi(x) = 0 \]
with
\[ \gamma^\nu \gamma^\mu \partial_\nu \partial_\mu = \frac{1}{2} \{\gamma^\nu, \gamma^\mu\} \partial_\nu \partial_\mu = \partial_\nu \partial_\mu \]
and so we get
\[ -(\partial_\nu \partial_\mu + m^2)\Psi^\alpha(x) = 0 \]
which is the Klein-Gordon equation. We show the spinor index \( \alpha \) explicitly to show that it applies to each component.
7.1.6 The slash

We will introduce some useful notation. Often we contract 4-vectors with gamma matrices and we write

$$A_\mu \gamma^\mu = \hat{A}$$

and the Dirac equation reads

$$(i\slashed{\partial} - m)\Psi = 0$$

7.1.7 Chiral spinors and the Weyl equation

We write our spinor in the Weyl basis as

$$\Psi(x) = \begin{pmatrix} \psi_L(x) \\ \psi_R(x) \end{pmatrix}$$

which is in the $\left(\frac{1}{2}, 0\right) \oplus (0, \frac{1}{2})$ representation. From XXX, we know that the Weyl spinors $\psi_L(x)$ in $(\frac{1}{2}, 0)$ and $\psi_R(x)$ in $(0, \frac{1}{2})$ transform the same way under rotations

$$S[A_{\text{rot}}] = \begin{pmatrix} \exp\left(\frac{i}{2} \theta \cdot \sigma\right) & 0_2 \\ 0_2 & \exp\left(-\frac{i}{2} \theta \cdot \sigma\right) \end{pmatrix}$$

(7.43)

but differently under boosts

$$S[A_{\text{boost}}] = \begin{pmatrix} \exp\left(\frac{1}{2} \eta \cdot \sigma\right) & 0_2 \\ 0_2 & \exp\left(-\frac{1}{2} \eta \cdot \sigma\right) \end{pmatrix}$$

(7.44)

or in components for rotations

$$\psi_{L/R} \rightarrow \exp\left(\frac{i}{2} \theta \cdot \sigma\right)\psi_{L/R}$$

and for boosts

$$\psi_{L/R} \rightarrow \exp(\pm \frac{1}{2} \eta \cdot \sigma)\psi_{L/R}$$

which decomposes the Dirac Lagrangian into

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

$$= i\psi_L^\dagger \sigma^\mu \partial_\mu \psi_L + i\psi_R^\dagger \sigma^\mu \partial_\mu \psi_R - m(\psi_L^\dagger \psi_R + \psi_R^\dagger \psi_L)$$

(7.45)

We see that a massive fermion seems to require both $\psi_L$ and $\psi_R$ since they couple through the mass term.\(^7\) In the massless case the two spinors decouple and we get the **Weyl equations**

$$i\sigma^\mu \partial_\mu \psi_L = 0$$

$$i\sigma^\mu \partial_\mu \psi_R = 0$$

We have already seen in the "derivation" of the Dirac equation in Sec. ??, that the equation of motion projects out half the degrees of freedom. The crucial aspect was, that the equation of motion was **first order** rather than second order as the Klein-Gordon equation.

\(^7\) You can also write a mass term with only one of them called a Majorana mass, see below.
In particular the conjugate momentum to the Dirac Lagrangian is given by

\[ \pi_\Psi = \frac{\partial L}{\partial \dot{\Psi}} = i\Psi^\dagger \]

which is not proportional to the time-derivative of \( \Psi \). Therefore, the phase space of the spinor is parametrized by \( \Psi \) and \( \Psi^\dagger \), while for a scalar it is \( \phi \) and \( \pi = \dot{\phi} \). So the phase space is described by 8 real dimensions and consequently the Dirac spinor has 4 real degrees of freedom.

### 7.1.8 \( \gamma^5 \)

In our choice of basis for the \( \gamma^\mu \) matrices, the Lorentz-transformations \( S[A] \) came out block-diagonal. What happens if we choose a different representation of the Clifford algebra?

\[ \gamma^\mu \rightarrow U\gamma^\mu U^{-1}, \quad \text{and} \quad \Psi \rightarrow U\Psi \]

\( S[A] \) will now likely not be block-diagonal. Can we define the Weyl spinors in an invariant way? We introduce a "fifth" gamma-matrix\(^8\)

\[ \gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \]

which satisfies which in the Weyl basis looks like

\[ \gamma^5 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \quad (7.46) \]

\[ \{\gamma^\mu, \gamma^5\} = 0, \quad \text{and} \quad (\gamma^5)^2 = +1 \]

As you will show in the exercises

\[ [S_{\mu\nu}, \gamma^5] = 0 \]

which means that \( \gamma^5 \) is a scalar under rotations and boosts. This is another manifestation of the fact, that the Dirac representation is reducible, since eigenvectors of \( \gamma^5 \) whose eigenvalues are different transform without mixing\(^9\).

We can therefore define Lorentz-invariant projection operators

\[ P_{L/R} = \frac{1}{2}(1 \mp \gamma^5) \]

such that

\[ P_L^2 = P_L, \quad P_R^2 = P_R, \quad P_LP_R = 0 \]

with Eq. (7.46), we see that \( P_{L/R} \) projects into the Weyl-basis

\[ \psi_L = P_L\Psi \]
\[ \psi_R = P_R\Psi \]

\( ^8 \) You can convince yourself that we can also write \( \gamma^5 \) as

\[ \gamma^5 = \frac{i}{4}\varepsilon^{\mu\nu\rho\sigma}\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma \]

using the totally anti-symmetric four index tensor with \( \varepsilon_{0123} = 1 \).

\( ^9 \) You can convince yourself that we can also write \( \gamma^5 \) as

\[ \gamma^5 = \frac{i}{4}\varepsilon^{\mu\nu\rho\sigma}\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma \]

using the totally anti-symmetric four index tensor with \( \varepsilon_{0123} = 1 \).

\( ^9 \) This criterium for reducibility is also known as Schur’s lemma.
7.2 Discrete symmetries

7.2.1 Parity

We already saw that under parity $P$ the left-handed and right-handed spinors are exchanged

$$ P : \psi_L \leftrightarrow \psi_R $$

Using this, we can write the action of parity on the Dirac spinor itself as

$$ P : \Psi(x, t) \rightarrow \gamma^0 \Psi(-x, t) $$

Notice that the Dirac Lagrangian conserves parity as you can easily see from Eq. (7.45). Let us now discuss how our interaction terms change under parity. The "mass" bi-linear transforms as

$$ P : \Psi(x, t) \Psi(x, t) \rightarrow \Psi(-x, t) \Psi(-x, t) $$

thanks to $(\gamma^0)^2 = 1$. This is the transformation of a scalar field. How about the vector $\Psi \gamma^\mu \Psi$? We split in temporal

$$ P : \Psi(x, t) \gamma^0 \Psi(x, t) \rightarrow \Psi(-x, t) \Psi(-x, t) $$

and spatial components

$$ P : \Psi(x, t) \gamma^i \Psi(x, t) \rightarrow \Psi(-x, t) \gamma^i \Psi(-x, t) $$

which shows that $\nabla \gamma^\mu \Psi$ is a vector. We can use $\gamma^5$ in the bilinear to form another Lorentz scalar and vector

$$ \nabla \gamma^5 \Psi, \quad \text{and} \quad \nabla \gamma^5 \gamma^\mu \Psi $$

We find with $\{\gamma^5, \gamma^\mu\} = 0$ that

$$ P : \nabla(x, t) \gamma^5 \Psi(x, t) \rightarrow \nabla(-x, t) \gamma^0 \gamma^5 \Psi(-x, t) $$

it is a pseudo-scalar and for the $\nabla \gamma^5 \gamma^\mu \Psi$ bilinear

$$ P : \nabla(x, t) \gamma^5 \gamma^\mu \Psi(x, t) \rightarrow \begin{cases} 
- \nabla(-x, t) \gamma^5 \gamma^0 \Psi(-x, t) \\
\nabla(-x, t) \gamma^5 \gamma^0 \Psi(-x, t) 
\end{cases} $$

we find that it is a pseudo-vector or axial-vector.

In summary we find the following bi-linears in Table 7.2.1 and their properties with

$$ \Psi(x) \Gamma \Psi(x) $$

for various $\Gamma$ matrices. The total number of bi-lines is 16, which is exactly what we could hope for from a $4 \times 4$ component object.

We have now a set of terms to build covariant Lagrangians with. If we add $\gamma^5$ to our terms then these terms will typically break

What about $\Gamma = \gamma^{[\mu \nu \sigma \tau]}$ or $\Gamma = \gamma^{[\mu \nu \sigma \tau \gamma]}$? The brackets indicate anti-symmetrization, since we do not need to consider the symmetric part which is $\sim \eta^{\mu\nu}$ (anti-symmetrization also explains why we consider at most 4-tensors). One can show that $\gamma^{[\mu \nu \sigma \tau]} = -i \epsilon^{\mu\nu\sigma\tau} \gamma_5$ and $\gamma^{[\mu \nu \sigma \tau \gamma]} = -i \epsilon^{\mu\nu\sigma\tau} \gamma_5$ and so they are linearly related to the simpler bilinears with $\Gamma = \gamma^\mu \gamma^5$ and $\Gamma = \gamma^5$. 

Like the rotation vector $J_i$ or the angular momentum $L = x \times p$. 

132 ANDREAS WEILER, TUM
parity. The electro-weak sector of the SM treats the $\psi_L$ and $\psi_R$ fields differently and they have different quantum numbers. They do not describe the same particle, e.g. the left-handed electron $e_L$ which couples to $W$-bosons of the $SU(2)_L$ gauge interactions is only distantly related to the right-handed electron $e_R$ which does not. A theory which treats $\psi_L$ and $\psi_R$ on the same footing is called a vector-like theory. A Lagrangian in which $\psi_L$ and $\psi_R$ appear differently, is called a chiral theory. QCD and QED are vector-like, and the electro-weak sector including the Higgs are chiral theories.

7.3 Solutions of the free Dirac equation

Let us study the solutions of the free Dirac equation

$$(\not{\gamma} - m)\Psi(p) = \begin{pmatrix} -m & \sigma^\mu p_\mu \\ \sigma^\mu p_\mu & -m \end{pmatrix} \Psi(p) = 0$$

(7.47)

We start with a simple ansatz

$$\Psi = u(p)e^{-ip\cdot x}, \quad \text{with} \quad p^2 = m^2$$

where the on-shell condition is motivated because the components satisfy the Klein-Gordon equation. For now we focus on positive frequencies, such that $p^0 > 0$. The column vector $u(p)$ needs to satisfy the additional constraint

$$(\gamma^\mu p_\mu - m)u(p) = 0$$

We can easily analyze this equation in the rest frame, where $p_r = (m, 0)$. We get the general solution by boosting with $S[\Lambda]$. In the rest frame

$$(m\gamma^0 - m)u(p_r) = m \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} u(p_r) = 0$$

the solutions are

$$u(p_r) = \sqrt{m} \begin{pmatrix} \xi \\ \xi \end{pmatrix}$$
for any two-component spinor $\xi$ which is just a number. We conventionally normalize

$$\xi^\dagger \xi = 1$$

and we added $\sqrt{m}$ for future convenience. The interpretation for the $\xi$ two spinor is given by the transformation under rotations, where $\xi$ transforms as an ordinary two-component spinor of the rotation group, see Eq. (7.43). We have

$$\xi \rightarrow \exp\left(\frac{i}{2} \theta \cdot \sigma\right)\xi$$

and $\xi$ therefore determines the spin orientation of the Dirac solution, as usual. E.g. when

$$\xi = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

then the particle has spin up (↑) along the $z$-direction. Notice again, that after applying the Dirac equation we are free to choose only 2 of the four components $u(p)$.

Let us now apply a boost in the $z$-direction

$$\begin{pmatrix} E \\ p^3 \end{pmatrix} = \begin{pmatrix} m \cosh \eta \\ m \sinh \eta \end{pmatrix}$$

Now apply the same boost to $u(p)$. With Eq. (7.44) we get

$$u(p) = \begin{pmatrix} \sqrt{p \cdot \sigma} \xi \\ \sqrt{p \cdot \bar{\sigma}} \xi \end{pmatrix}$$

(7.48)

with $p = (E, 0, 0, p^3)$ and where the square-root of a matrix is understood as taking the positive root of each eigenvalue. Let us show that this satisfies the Dirac equation. With $u(p)^T = (u_1, u_2)$ we can write the Dirac equation Eq. (7.47) as

$$\begin{align*}
(p \cdot \sigma) u_2 &= m u_1 \\
(p \cdot \bar{\sigma}) u_1 &= m u_2
\end{align*}$$

(7.49)  (7.50)

We can see that either of the two equations implies the other using

$$(p \cdot \sigma)(p \cdot \bar{\sigma}) = p_0^2 - p_i p_j \sigma^i \sigma^j = p_0^2 - p_i p_j \delta^{ij} = p_\mu p^\mu = m^2$$

Let us try the ansatz $u_1 = p \cdot \sigma \xi'$ for some $\xi'$ then Eq. (7.50) immediately gives

$$u_2 = m \xi'$$

We find that any spinor of the form

$$u(p) = A \begin{pmatrix} (p \cdot \sigma) \xi' \\ m \xi' \end{pmatrix}$$

with constant $A$ would be solution. To make this more symmetric, we choose $A = 1/m$ and $\xi' = \sqrt{p \cdot \bar{\sigma}} \xi$ with constant $\xi$. Then $u_1 = (p \cdot \sigma) \sqrt{p \cdot \bar{\sigma}} = m \sqrt{p \cdot \bar{\sigma}} \xi$
7.3.1 Negative energy solutions

We get another solution to the Dirac equation with the ansatz
\[ \Psi = v(p) e^{ip \cdot x} \]

The \( u(p) \) solution oscillates in time as \( \Psi \sim e^{-iEt} \) and is called a **positive frequency** solution. The \( v(p) \) solutions oscillates with \( \Psi \sim e^{+iEt} \) and are called **negative frequency** solutions. Both are solutions to the classical Dirac equation and note that both have **positive energy**. The Dirac equation requires that the 4-component spinor \( v(p) \) satisfies
\[
\begin{pmatrix}
    m & \sigma^\mu p_\mu \\
    \sigma^\mu p_\mu & m
\end{pmatrix} v(p) = 0
\]

which with the above we see is solved by
\[
v(p) = \begin{pmatrix}
    \sqrt{p \cdot \sigma} \chi \\
    -\sqrt{p \cdot \sigma} \chi
\end{pmatrix}
\]

with some normalized \( (\chi^\dagger \chi = 1) \) 2-component spinor, taken to be a constant.

7.3.2 Examples

We take again the positive frequency solution with mass \( m \) at rest \( p = 0 \)
\[ u(p_r) = \sqrt{m} \begin{pmatrix}
    \xi \\
    \xi
\end{pmatrix} \]

Consider \( \xi^T = (1, 0) \) we boost to \((E, 0, 0, p^3)\) to get
\[
u(p) = \begin{pmatrix}
    \sqrt{p \cdot \sigma} \begin{pmatrix}
        1 \\
        0
\end{pmatrix} \\
    \sqrt{p \cdot \sigma} \begin{pmatrix}
        0 \\
        1
\end{pmatrix}
\end{pmatrix} = \begin{pmatrix}
    \sqrt{E - p^3} \begin{pmatrix}
        1 \\
        0
\end{pmatrix} \\
    \sqrt{E + p^3} \begin{pmatrix}
        1 \\
        0
\end{pmatrix}
\end{pmatrix}
\]

This expression in fact also makes sense for a massless particle with \( E = p^3 \). For a massless spinor we therefore obtain
\[ u(p) = \sqrt{2E} \begin{pmatrix}
    0 \\
    0 \\
    1
\end{pmatrix} \]

Similarly, if we boost \( \xi^T = (0, 1) \) (spin down), we get
\[
u(p) = \begin{pmatrix}
    \sqrt{p \cdot \sigma} \begin{pmatrix}
        0 \\
        1
\end{pmatrix} \\
    \sqrt{p \cdot \sigma} \begin{pmatrix}
        0 \\
        1
\end{pmatrix}
\end{pmatrix} \xrightarrow{m \to 0} \sqrt{2E} \begin{pmatrix}
    0 \\
    1 \\
    0
\end{pmatrix}
\]

We picked the normalization such that this limit would make sense.
7.3.3 Helicity

We finally explain the meaning of "left-handed" for $\psi_L$ and "right-handed" for $\psi_R$. The massless solutions are eigenstates of the helicity operator

$$h \equiv \hat{p} \cdot S = \frac{1}{2} \hat{p}_i \sigma^i = \frac{1}{2} \frac{\sigma}{|\sigma|} \begin{pmatrix} \sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix}$$ (7.52)

where $\hat{p}_i = p_i/|p|$. A particle with $h = +1/2$ is called right-handed, while one with $h = -1/2$ is called left-handed.

The helicity of a massive particle depends on the reference frame: we can always boost to a frame in which its momentum is in the opposite direction (but its spin stays unchanged). 10

The Lorentz-invariance of helicity (for mass-less particles) is manifest in the representation of Weyl spinors $\psi_L$ and $\psi_R$ which live in different representations of the Lorentz group.

7.3.4 Useful results: inner and outer products

It is convenient to introduce a basis $\xi^s$ and $\chi^s$ with $s = 1, 2$ for the two-component spinores such that

$$\xi^r \xi^s = \delta^{rs}, \text{ and } \chi^r \chi^s = \delta^{rs}$$

which e.g. is satisfied for

$$\xi^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \text{ and } \xi^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and similarly for $\chi^s$. The two independent solutions to the positive frequency ansatz are

$$u^s(p) = \begin{pmatrix} \sqrt{p \cdot \sigma} \xi^s \\ \sqrt{p \cdot \bar{\sigma}} \xi^s \end{pmatrix}$$

We can take the inner product in two ways, either $u^\dagger \cdot u$ or $\bar{u} \cdot u$. Only $\bar{u} \cdot u$ will be Lorentz-invariant of course, but we will need $u^\dagger \cdot u$ to quantize the theory later. Let us determine the results.

$$(u^r(p))^\dagger \cdot u^s(p) = (\xi^{r\dagger} \sqrt{p \cdot \sigma}, \xi^{r\dagger} \sqrt{p \cdot \bar{\sigma}}) \begin{pmatrix} \sqrt{p \cdot \sigma} \xi^s \\ \sqrt{p \cdot \bar{\sigma}} \xi^s \end{pmatrix}$$

$$= \xi^{r\dagger} p \cdot \sigma \xi^s + \xi^{r\dagger} p \cdot \bar{\sigma} \xi^s$$

$$= 2 \xi^{r\dagger} p_0 \xi^s$$

$$= 2E \delta^{rs}$$

and for the Lorentz-invariant inner product

$$\bar{u}^r(p) \cdot u^s(p) = (\xi^{r\dagger} \sqrt{p \cdot \sigma}, \xi^{r\dagger} \sqrt{p \cdot \bar{\sigma}}) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \sqrt{p \cdot \sigma} \xi^s \\ \sqrt{p \cdot \bar{\sigma}} \xi^s \end{pmatrix} = 2m \delta^{rs}$$

In Ex9, you will discuss the Pauli-Lubanski polarization vector,

$$W_\mu \equiv (1/2)\epsilon_\mu\nu\sigma^\rho p^\nu J^\rho$$

which is a Lorentz covariant operator. Using the commutation relations for the generators of the Poincaré group, one can show that $P^2 = P\mu P^\mu$ and $W^2 = W_\mu W^\mu$, are Casimir operators for the Poincaré algebra. Since the Poincaré group has rank 2, $P^2$ and $W^2$ are its only Casimir operators. Therefore a massive state can be labeled by two numbers, its mass and spin. A massless state can be labeled by only one number (called helicity), since $P^\mu$ and $W^\mu$ for massless particles are proportional to each other.

10 For a massless particle moving with $v = c$, we cannot perform such a boost.
Similarly for the negative frequency solutions we have
\[ v^s(p) = \begin{pmatrix} \sqrt{p \cdot \sigma} \chi^s \\ -\sqrt{p \cdot \sigma} \chi^s \end{pmatrix} \]
with
\[
(v^r(p))^\dagger \cdot v^s(p) = 2p_0 \delta^{rs} \\
v^r(p) \cdot v^s(p) = -2m \delta^{rs}
\]
What about the inner products of \( u \) and \( v \)? We compute
\[
\bar{u}^r(p) \cdot v^s(p) = (\xi^r \sqrt{p \cdot \sigma} \chi^s) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \sqrt{p \cdot \sigma} \chi^s \\ -\sqrt{p \cdot \sigma} \chi^s \end{pmatrix} = 0
\]
and similarly
\[
\bar{v}^r(p) \cdot u^s(p) = 0
\]
For the adjoint inner product, there is another combination which has useful properties. With \((p')^\mu = (p^0, -\vec{p})\), we have
\[
(u^r(p))^\dagger \cdot v^s(p') = (\xi^r \sqrt{p' \cdot \sigma} \chi^s) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \sqrt{p' \cdot \sigma} \chi^s \\ -\sqrt{p' \cdot \sigma} \chi^s \end{pmatrix} (7.53)
\]
\[
= \xi^r (\sqrt{p \cdot \sigma} p' \cdot \sigma) \chi^s + \xi^r \sqrt{p \cdot \sigma} p' \cdot \sigma \chi^s = 0
\]
The terms under the square root are
\[
(p \cdot \sigma)(p' \cdot \sigma) = (p_0 + p_i \sigma^i)(p_0 - p_i \sigma^i) = p_0^2 - \vec{p}^2 = m^2
\]
and the same holds for \((p \cdot \bar{\sigma})(p' \cdot \bar{\sigma})\). This means the two terms in Eq. (7.54) cancel and we get
\[
(u^r(p))^\dagger \cdot v^s(-\vec{p}) = (v^r(p))^\dagger \cdot u^s(-\vec{p}) = 0
\]
\[7.3.5 \quad \text{Outer products}\]
A final spinor identity will prove very useful before we turn to quantizing everything. We claim
\[
\sum_{s=1}^{2} u^s(p) \bar{u}^s(p) = \hat{\rho} + m
\]
Note, the two spinors are not contracted, but placed back to back to give a 4 \times 4 matrix Similarly
\[
\sum_{s=1}^{2} v^s(p) \bar{v}^s(p) = \hat{\rho} - m
\]
Now let us show this.
\[
\sum_{s=1}^{2} u^s(p) \bar{u}^s(p) = \sum_{s=1}^{2} \begin{pmatrix} \sqrt{p \cdot \sigma} \xi^s \\ \sqrt{p \cdot \bar{\sigma}} \xi^s \end{pmatrix} \begin{pmatrix} \xi^s \sqrt{p \cdot \sigma} \\ \xi^s \sqrt{p \cdot \bar{\sigma}} \end{pmatrix} = 0
\]
Now

\[
\sum_{s=1}^{2} \xi^s \xi^{s\dagger} = 1_2
\]

which gives for Eq. (7.56)

\[
\sum_{s=1}^{2} u^s(p) \bar{u}^s(p) = \begin{pmatrix} m & p \cdot \sigma \\ p \cdot \bar{\sigma} & m \end{pmatrix} = \not{p} + m \quad (7.57)
\]

You can now easily prove \(\sum_{s=1}^{2} v^s(p) \bar{v}^s(p) = \not{p} - m\).

---

**The Dirac Sea:**

A historical detour: if we interpret \(\Psi\) as a single-particle wave function, the plane wave solutions to the Dirac equation are energy eigenstates, with

\[
\Psi = u(p)e^{-ip \cdot x} \sim i \frac{\partial \Psi}{\partial t} = +E_p \Psi \\
\Psi = v(p)e^{ip \cdot x} \sim i \frac{\partial \Psi}{\partial t} = -E_p \Psi
\]

which look like positive and negative energy solutions. The spectrum is unbounded below! This seems disastrous. Dirac invented the 'Dirac sea', where all negative energy states are filled up thanks to the Pauli-exclusion principle. He postulated only charge differences are observable (similar to the energy difference argument motivating normal ordering). He postulated that if you excite a negative energy state, it leaves a hole with all the properties of an electron, except that it carries positive charge! This is the prediction of anti-matter.

We now understand that the interpretation of the Dirac spinor as a single-particle wavefunction is not really correct. The picture of a Dirac sea relies on the fact, that the particles are fermions – which breaks down for a charged scalar, a boson, which also predicts anti-particles.

What we can still learn from Dirac’s analysis is that there is no consistent way to interpret the Dirac equation as describing a single particle. We need to think of it as a classical field equation with only positive energy solutions (the Hamiltonian is positive definite!). Once we quantize the theory, we encounter particle and anti-particle excitations.

---

### 7.4 Symmetries of the Dirac Lagrangian

We can proceed as in the scalar Lagrangian and derive the spacetime symmetries

\[
\delta \Psi = \varepsilon^\mu \partial_\mu \Psi
\]

and use the standard formula Eq. (4.67) for the energy-momentum tensor

\[
T^{\mu\nu} = i \bar{\Psi} \gamma^\mu \partial^\nu \Psi - \mathcal{L}
\]

using the fact that \(\mathcal{L}\) does not depend on \(\partial_\mu \bar{\Psi}\). Since the current is only conserved on-shell, we can impose the equations of motions directly on \(T^{\mu\nu}\) and we won’t lose anything. Since the equations of motions are first order, this affects the energy-momentum tensor. We use \((i\gamma^\mu \partial_\mu - m)\Psi(x) = 0\) which means, we can set \(\mathcal{L} = 0\), to obtain

\[
T^{\mu\nu} = i \bar{\Psi} \gamma^\mu \partial^\nu \Psi
\]
which e.g. leaves for the total energy

\[ E = \int d^3x T^{00} = \int d^3x i\overline{\Psi}\gamma^0\Psi \]  

(7.58)

\[ = \int d^3x i\overline{\Psi}i\gamma^0(-i\gamma^\mu\partial_\mu + m)\Psi \]  

(7.59)

In the last equality we have again used the equations of motion.

### 7.4.1 Internal Symmetries

We can rotate the Dirac spinor by a phase

\[ \Psi \rightarrow e^{i\alpha}\Psi \]

which leaves the Dirac Lagrangian invariant. This gives rise to a current

\[ j^\mu_V = \overline{\Psi}\gamma^\mu\Psi \]  

(7.60)

where we specify that it is a vector symmetry which transforms \(\psi_L\) and \(\psi_R\) in the same way. We can easily obtain \(j^\mu_V\) using the Noether procedure or we can simply check that it is conserved on-shell:

\[ \partial_\mu j^\mu_V = (\partial_\mu \overline{\Psi})\gamma^\mu\Psi + \overline{\Psi}\gamma^\mu(\partial_\mu\Psi) \]

\[ = im\overline{\Psi}\psi - im\overline{\Psi} = 0 \]

where we have used \(i\partial_\Psi = m\overline{\Psi}\) and the Dirac adjoint equation \(\overline{\Psi}i\slashed{\partial} = -m\overline{\Psi}\), where the arrow points to the term we take the derivative of. The conserved charge

\[ Q = \int d^3x \overline{\Psi}\gamma^0\Psi = \int d^3x \overline{\Psi}\gamma^0\Psi \]  

(7.61)

has the interpretation of electric charge or particle number for fermions.

For a massless Dirac spinor \(m = 0\), the left-handed \(\psi_L\) and right-handed \(\psi_R\) spinors decouple and we can rotate the left and right-handed fermions in opposite directions

\[ \Psi \rightarrow e^{\pm i\alpha}\gamma^5 \]  

(7.62)

which gives the conserved current

\[ j^\mu_A = \overline{\Psi}\gamma^5\gamma^\mu\Psi \]  

(7.63)

This is only conserved when \(m = 0\), as we can see by taking the divergence

\[ \partial_\mu j^\mu_A = (\partial_\mu \overline{\Psi})\gamma^\mu\gamma^5\Psi + \overline{\Psi}\gamma^\mu\gamma^5(\partial_\mu\Psi) \]

\[ = (\partial_\mu \overline{\Psi})\gamma^\mu\gamma^5\Psi - \overline{\Psi}\gamma^5\gamma^\mu(\partial_\mu\Psi) \]

\[ = 2im\overline{\Psi}\gamma^5\Psi \]

which vanishes only for \(m = 0\). Once we couple an axial current to a gauge field (similar to the photon field \(A_\mu\)), axial symmetries do not survive the quantization process, they develop anomalies. An Anomaly is a symmetry of the classical theory which is not preserved in the quantum theory.\(^\text{11}\)

\(^{11}\) Another symmetry which is usually anomalous is scale symmetry \(x^\mu \rightarrow e^{-\alpha x^\mu}\), see Ex4-2.
7.5 Quantizing the Dirac Field

If we were to use the same approach for the free Dirac Lagrangian
\[ \mathcal{L} = \overline{\Psi}(x)(i\gamma^\mu \partial_\mu - m)\Psi(x) \]
as for the Klein-Gordon scalar theory, we would run into serious trouble: the quantized free Hamiltonian would seem to be able to lower the energy by creating particles (The Hamiltonian is not bounded below). If we tried anti-commutation relations for a scalar field, the cancellation between negative and positive frequencies to guarantee causality in Eq. (5.71) and Eq. (5.72) would not work anymore. This is the gist of the spin-statistics theorem.

7.5.1 The Hamiltonian

We will need the Hamiltonian. With the conjugate momentum
\[ \pi = i\Psi^\dagger \]
we obtain
\[ \mathcal{H} = \pi \dot{\Psi} - \mathcal{L} = \overline{\Psi}(-i\gamma^i \partial_i + m)\Psi \tag{7.64} \]
which is the same as the result we obtained deriving the energy-momentum tensor in Eq. (7.59).

We now want to turn this into an operator. We will expand \( \Psi(x) \) into a basis of eigenfunctions of the free Dirac equation
\[ \psi(x) = \sum_{s=1}^{2} \int \frac{d^3p}{\sqrt{2E_p}} \left( b_p^s u^s(p)e^{+ip\cdot x} + c_p^s v^s(p)e^{-ip\cdot x} \right) \]
\[ \psi^\dagger(x) = \sum_{s=1}^{2} \int \frac{d^3p}{\sqrt{2E_p}} \left( b_p^s \dagger u^s(p)e^{-ip\cdot x} + c_p^s \dagger v^s(p)e^{+ip\cdot x} \right) \]
where the operators \( b_p^s \dagger \) create particles associated with the spinor state \( u^s(p) \), while the \( c_p^s \dagger \) create anti-particles associated to \( v^s(p) \). Except for the \( u^s(p) \) and \( v^s(p) \), this has the exact same form as the expansion of a complex scalar field, see Sec. 5.5.

7.5.2 Quantization of fermions

We need to account for the Fermi-Dirac-statistics for fermions, meaning that the state picks up a minus sign upon the interchange of any two particles. The spin-statistics theorem says that integer spin must be quantized as bosons, while half-integer spin must be quantized as fermions. If we try to do otherwise (as you should in the exercise), we will find inconsistencies, such as unboundedness of the free Hamiltonian.

How do we quantize fermionic fields? Recall for bosons
\[ [a_p^\dagger, a_q^\dagger] = 0 \quad \sim \quad a_p^\dagger a_q^\dagger |0\rangle \equiv |p, q\rangle = |q, p\rangle \]
If we want Fermi-Dirac statistics, we will need anti-commutation relations \( \{A, B\} = AB + BA \) to hold.
Recall that for a scalar field $\phi$, we could interpret $a_p^\dagger$ and $a_p$ as creation and annihilation operators because of their commutation relation with the number operator:\[ N_\phi = \int \frac{d^3k}{(2\pi)^3} a_k^\dagger a_k \]

Recall, that this worked because of the following identity for commutators

\[ [AB,C] = A[B,C] + [A,C]B \]

which immediately gives\[ \left[ N, a_{k'}^\dagger \right] = \int \frac{d^3k}{(2\pi)^3} \left[ a_{k}a_{k'}, a_{k'}^\dagger \right] = a_{k'}^\dagger. \]

Therefore $a_{k'}^\dagger$, acting on a state raises the eigenvalue of $N$ by one and the energy by $E_k$, as expected for a creation operator. However there is another useful relation for commutators

\[ [AB,C] = A\{B,C\} - \{A,C\}B \]

with the anticommutator $\{A,B\} = AB + BA$. This means that if we were to impose anticommutation relations on the $a_{k'}^\dagger$ and $a_k$, then they could still have an interpretation as creation an annihilation operators!

We will therefore ask the spinor field to satisfy

\[ \{\Psi_\alpha(x), \Psi_\beta(y)\} = \{\Psi_\alpha^\dagger(x), \Psi_\beta^\dagger(y)\} = 0 \]

and

\[ \left\{ \Psi_\alpha(x), \Psi_\beta^\dagger(y) \right\} = \delta_{\alpha\beta}\delta^{(3)}(x-y) \]

Using the expansion for $\Psi$ and $\Psi^\dagger$ gives

\[ \{b_p^\dagger, b_q^\dagger\} = \{c_p^\dagger, c_q^\dagger\} = \{b_p^\dagger, c_q^\dagger\} = \ldots = 0 \]

and the canonical non-zero relations

\[ \{b_p^\dagger, b_q^\dagger\} = (2\pi)^3 \delta^{(3)}(p-q) \]
\[ \{c_p^\dagger, c_q^\dagger\} = (2\pi)^3 \delta^{(3)}(p-q) \]

The calculation of the Hamiltonian is straight-forward but somewhat tedious and we get

\[ H = \int \frac{d^3p}{(2\pi)^3} E_p \left( b_p^\dagger b_p^\dagger - c_p^\dagger c_p^\dagger \right) \]
\[ = \int \frac{d^3p}{(2\pi)^3} E_p \left( b_p^\dagger b_p^\dagger + c_p^\dagger c_p^\dagger - (2\pi)^3\delta^{(3)}(0) \right) \]

where the anti-commutation relation saved us from the indignity of an unbounded Hamiltonian! Interestingly the contribution to the cosmological constant now has the opposite sign compared to the bosonic result.\[^{14}\]

\[^{12}\] or equivalently with the free Hamiltonian

\[ H = \int \frac{d^3k}{(2\pi)^3} E_k a_k^\dagger a_k \]

\[^{13}\] and also

\[ [N, a_{k'}] = -a_{k'} \]

Recall that this means for eigenstates of the number operator

\[ N|n\rangle = n|n\rangle \]

and so

\[ N a_{k'}^\dagger|n\rangle = ([N, a_{k'}^\dagger] + a_{k'}^\dagger N)|n\rangle = (a_{k'}^\dagger + a_{k'} N)|n\rangle = (1 + n)a_{k'}^\dagger|n\rangle \]

and so

\[ a_{k'}^\dagger|n\rangle \propto |n + 1\rangle \]

up to normalization.

\[^{14}\] This anticipates a result of supersymmetry, which is a symmetry between fermions and bosons. Exact supersymmetry predicts $H(0) = ((2\pi)^3\delta^{(3)}(0) - (2\pi)^3\delta^{(3)}(0))(0) = 0$. This holds after quantization and even in the presence of interactions!
7.5.3 Fermi-Dirac Statistics

We define the (free) vacuum as usual

$$b_p^* |0\rangle = c_p^* |0\rangle = 0$$

for all creation operators. The Hamiltonian still has nice commutation relations with the anti-commuting $b$ and $c$ operators. You can easily see with Eq. (7.65) that

$$[H, b_{p_1}^\dagger] = E_{p_1} b_{p_1}^\dagger$$
$$[H, b_{p_1}^\dagger] = -E_{p_1} b_{p_1}^\dagger$$

and exactly the same for $b \leftrightarrow c$. We can treat them therefore again as ladder operators creating a tower of energy-eigenstates by acting with $b_{p_1}^\dagger$ and $c_{p_1}^\dagger$ on the vacuum, just as in the bosonic case. The (not relativistically normalized) one-particle states are

$$|p, r\rangle = b_{p_1}^\dagger |0\rangle$$

and the (not relativistically normalized) two-particle states satisfy

$$|p_1, r_1; p_2, r_2\rangle = (b_{p_1}^\dagger)^i (b_{p_2}^\dagger)^j |0\rangle = -|p_2, r_2; p_1, r_1\rangle$$

exactly as we wanted, confirming that the particles do indeed obey Fermi-Dirac statistics. In particular, we obtain the **Pauli-Exclusion principle**

$$|p_1, r_1; p_1, r_1\rangle = 0$$

As for the scalar field (which had spin 0), we could act with the angular momentum operator and confirm that a particle at rest

$$|p = 0, r\rangle$$

does indeed carry intrinsic angular momentum $1/2$ as expected.

7.5.4 Causal Fermi Fields

It is not clear that a theory with Fermi-Dirac statistics will lead to causal predictions. For bosons, we said that

$$[\phi(x), \phi(y)] = 0, \text{ for } (x - y)^2 < 0.$$ 

For Fermi fields we now have the relations

$$\{\Psi_\alpha(x), \Psi_\beta(y)\} = 0$$

and

$$\{\Psi_\alpha(x), \Psi_\beta^\dagger(y)\} = 0, \text{ for } (x - y)^2 < 0$$

How do we see that this guarantees causality in the quantized theory?
The reason is that observables are necessarily always bilinear in the fields, e.g. the energy Eq. (??), momentum and fermion number Eq. (7.61) (or charge) are given by

\[ E = \int d^3x \, i \Psi^\dagger \partial_t \Psi \]

\[ P^i = \int d^3x \, i \Psi^\dagger \partial^i \Psi \]

\[ Q = \int d^3x \, \Psi^\dagger \Psi \]

This is not entirely surprising: we know that spinors form a double-valued representation of the Lorentz group since they change sign under $2\pi$ rotations. Observables, however, are unaffected by a rotation by $2\pi$ and so must be composed of an even number of spinor fields.

### 7.6 Perturbation Theory for Spinors

We will consider a simple nucleon-meson theory\(^{15}\)

\[ \mathcal{L} = \bar{\Psi} (i \gamma^\mu \partial_\mu - m) \Psi + \frac{1}{2} (\partial_\mu \phi)^2 - \frac{\mu^2}{2} \phi^2 - \lambda \bar{\Psi} \Gamma \Psi \phi \]  

(7.70)

with $\Gamma = 1$ in case of $\phi$ being a scalar or $\Gamma = i \gamma^5$ for a pseudo-scalar.\(^{16}\)

Note, the dimensions of the fields, as before $[\phi] = 1$, but the kinetic term of the spinor requires $[\Psi] = 3/2$. This turns the interaction to a marginal term $[\lambda] = 0$, compare to the dimensionful (relevant) interaction in the case of a trilinear coupling between scalars.

Dyson’s formula and Wick’s theorem can be used in almost the same way for Fermi fields as for scalars. However, the anticommutation relations lead to a crucial difference. Remember that for space-like distances $(x - y)^2 < 0$, time-ordering is not Lorentz-invariant.\(^{17}\) Still, the $T$ product of two scalar fields is Lorentz-invariant because the fields commute for space-like distances $(x - y)^2 < 0$, so

\[ \phi(x)\phi(y) = \phi(y)\phi(x) \]

and the order here doesn’t matter. For fermions this no longer holds! If $(x - y)^2 < 0$, Fermi fields anticommute. For $(x - y)^2 < 0$ and if we are in a frame where $x^0 > y^0$

\[ T\{\Psi(x)\Psi(y)\} = \Psi(x)\Psi(y) \]  

(7.71)

in the Lorentz-transformed frame where $y^0 > x^0$, we have

\[ T\{\Psi(x)\Psi(y)\} = \Psi(y)\Psi(x) = -\Psi(x)\Psi(y) \]  

(7.72)

We conclude that we must modify the definition of the $T$ ordering of Fermi-fields to make it Lorentz invariant.

---

\(^{15}\) The nucleons are fermions now and we are graduating from a toy model to something closer to Nature, even though it does not in fact describe nucleon-meson interactions at low energy correctly.

\(^{16}\) We need the $i$ for hermiticity of $\mathcal{L}$.

\(^{17}\) Since we can boost to a reference frame where the temporal order $x^0 > y^0$ is reversed to $x^0 < y^0$. 
The solution is straight-forward. We define the $T$-product to include a factor of $(-1)^n$ where $n$ is the number of exchanges required to time order the Fermi fields.

For two fields

$$T\{\Psi(x)\Psi(y)\} = \begin{cases} 
\Psi(x)\Psi(y) & \text{for } x^0 > y^0 \\
-\Psi(y)\Psi(x) & \text{for } y^0 > x^0 
\end{cases} \quad (7.73)$$

in particular we have

$$T\{\Psi_1(x)\Psi_2(y)\} = -T\{\Psi_2(y)\Psi_1(x)\}$$

The normal-ordering product is defined as before and we define with

$$\Psi = \Psi^+ + \Psi^-$$

where $\Psi^+$ multiplies an annihilation operator, and $\Psi^-$ multiplies a creation operator.

$$: \Psi_1 \Psi_2 : = : \Psi_1^+ \Psi_2^+ + \Psi_1^- \Psi_2^- + \Psi_1^+ \Psi_2^- + \Psi_1^- \Psi_2^+ :$$

$$= \Psi_1^+ \Psi_2^+ - \Psi_2^- \Psi_1^- + \Psi_1^- \Psi_2^+ + \Psi_1^+ \Psi_2^-$$

where the second picked up a $(-1)$ because we had to interchange to Fermi fields. Just as for the $T$ product, we therefore have

$$: \Psi_1 \Psi_2 : = - : \Psi_2 \Psi_1 :$$

With these modifications, we can use Dyson’s formula and Wick’s theorem as before. We just need to be careful with contractions in Wick’s theorem, e.g.

$$: \Psi_1 \Psi_2 \Psi_3 \Psi_4 : = - : \Psi_1 \Psi_3 \Psi_2 \Psi_4 : = - \Psi_1 \Psi_3 : \Psi_2 \Psi_4 :$$

### 7.7 The Fermion Propagator

We will now move to the Heisenberg picture and define spinor operators at every point in spacetime $\Psi(x, t)$ such that they evolve according to

$$\frac{\partial \Psi}{\partial t} = i[H, \Psi]$$

which is the usual operator time evolution. As before, we can solve this by expanding (replacing $i \mathbf{x} \cdot \mathbf{p} \rightarrow -i \mathbf{x} \cdot \mathbf{p}$ in the exponent)

$$\Psi(x) = \sum_{s=1}^{2} \frac{d^3 p}{\sqrt{2E_p}} \left( b_p^s u^s(p)e^{-ip \cdot x} + c_p^s u^s(p)e^{ip \cdot x} \right) \quad (7.74)$$

$$\Psi^\dagger(x) = \sum_{s=1}^{2} \frac{d^3 p}{\sqrt{2E_p}} \left( b_p^s u^s(p)e^{ip \cdot x} + c_p^s u^s(p)e^{-ip \cdot x} \right) \quad (7.75)$$

Remember the strange historic convention stemming from the positive/negative energy in the exponent.

Recall that for scalar fields, their order in $T$ and $N$ products was unimportant.
We define the fermionic propagator to be
\[ iS_{\alpha\beta} \equiv \{ \Psi_\alpha, \overline{\Psi}_\beta \} \tag{7.76} \]

Inserting the expansion Eq. (7.74) we obtain
\[
iS_{\alpha\beta}(x - y) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2Ep} \left[ \{ b^\dagger_\alpha(p), b_\beta(q) \} e^{-ip \cdot (x - y)} + \{ \epsilon_\alpha(p), \epsilon^*_\beta(q) \} e^{+ip \cdot (x - y)} \right]
\]

We use Eq. (7.55) and the anticommutation relations for \( b \) and \( c \) of Eq. (7.68) to obtain
\[
iS_{\alpha\beta}(x - y) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2Ep} \left[ (\hat{p} + m)_{\alpha\beta} e^{-ip \cdot (x - y)} + (\hat{p} - m)_{\alpha\beta} e^{+ip \cdot (x - y)} \right]
\]

We can therefore write
\[
iS(x - y) = (i\partial_x + m)(D(x - y) - D(y - x))
\]
in terms of the propagator for the real scalar field \( D(x - y) \). If we stay away from singularities, the propagator satisfies
\[
(i\partial_x - m)S(x - y) = 0
\]
which directly follows because
\[
(i\partial_x - m)S(x - y) = (i\partial_x - m)(i\partial_x + m)(D(x - y) - D(y - x)) = - (\Box + m^2)(D(x - y) - D(y - x)) = 0
\]
using the properties of the scalar propagator for \( p^2 = m^2 \) away from the singularities.

### 7.7.1 The Feynman propagator for Dirac fields

Similarly, we obtain first for the vacuum expectation value
\[
\langle 0 | \Psi_\alpha(x) \overline{\Psi}_\beta(y) | 0 \rangle = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2Ep} (\hat{p} + m)_{\alpha\beta} e^{-ip \cdot (x - y)}
\]
\[
= (i\partial_x + m)D(x - y)
\]
and
\[
\langle 0 | \overline{\Psi}_\beta(y) \Psi_\alpha(x) | 0 \rangle = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2Ep} (\hat{p} - m)_{\alpha\beta} e^{+ip \cdot (x - y)}
\]
\[
= (-i\partial_x - m)D(y - x)
\]
from which we define the Feynman propagator
\[
S_F(x - y) = \overline{\Psi(x)} \Gamma \Psi(y)
\]
We will in the following often drop the indices, but you should remember that \( S(x - y) = \{ \Psi(x), \overline{\Psi}(y) \} \) is a \( 4 \times 4 \) matrix. Recall, we can write
\[
D(x - y) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2Ep} e^{-ip \cdot (x - y)}
\]
with \( p^\mu = (Ep, \mathbf{p}) \) or \( p^2 = m^2 \). Note, it is also a \( 4 \times 4 \) matrix.
We are interested in which satisfies the initial state is for the adjoint we therefore have this process in Sec. 6.4.1, except that the nucleons now have spin. Let us study \( \Psi \Psi \rightarrow 7.7.2 \text{ Nucleon-Nucleon scattering with spin} \)

We conclude that \( S \) thanks to \( \phi(x)\phi(y) = \Delta_F(x - y) = \int \frac{d^4p}{(2\pi)^4} e^{-i p \cdot (x-y)} \frac{i}{p^2 - m^2 + i\varepsilon} \)

The integral representation of the Feynman propagator for Dirac spinors is therefore

\[
S_F(x - y) = i \int \frac{d^4p}{(2\pi)^4} e^{-i p \cdot (x-y)} \frac{\not{p} + m}{p^2 - m^2 + i\varepsilon} \tag{7.77}
\]

which satisfies

\[
(i\slashed{\partial}_x - m)S_F(x - y) = i\delta^{(4)}(x - y)
\]
	hanks{\( (i\slashed{\partial}_x - m)S_F(x - y) = i\delta^{(4)}(x - y) \) derived in Eq. (5.79). We conclude that \( S_F(x - y) \) is a Green’s function of the Dirac operator.}

7.7.2 Nucleon-Nucleon scattering with spin

Let us study \( \Psi \Psi \rightarrow \Psi \Psi \) scattering again. We have already discussed this process in Sec. 6.4.1, except that the nucleons now have spin.

The initial state is

\[
|i\rangle = |p_1, s; p_2, r\rangle = \sqrt{2E_{p_1}} \sqrt{2E_{p_2}} b_{p_1}^\dagger b_{p_2}^\dagger |0\rangle \tag{7.78}
\]

and the final state is

\[
|f\rangle = |q_1, s'; q_2, r'\rangle = \sqrt{2E_{q_1}} \sqrt{2E_{q_2}} b_{q_1}^\dagger b_{q_2}^\dagger |0\rangle \tag{7.79}
\]

We need to be careful with signs, since the \( b^\dagger \) now anti-commute.

For the adjoint we therefore have

\[
\langle f | = \sqrt{4E_{q_1}E_{q_2}} \langle 0 | b_{q_2}^\dagger b_{q_1}^\dagger \tag{7.80}
\]

We are interested in \( \langle f |S - 1|i\rangle \) at order \( \lambda^2 \),

\[
\frac{(-i\lambda)^2}{2} \int d^4x_1 d^4x_2 T \left[ \overline{\Psi}(x_1)\Psi(x_1)\phi(x_1)\overline{\Psi}(x_2)\Psi(x_2)\phi(x_2) \right] \tag{7.81}
\]
where as usual, all field operators are in the interraction picture. As in the pure-scalar calculation, the contribution to Nucleon-Nucleon scattering is due to the contraction

\[
\phi(x_1) \phi(x_2) \Psi(x_1) \bar{\Psi}(x_2) \Psi(x_2) \bar{\Psi}(x_1) \theta(x_1 - x_2)
\]

We now have to be careful about how the spinor indices are contracted. Let us start with the \(\Psi\) and deal with \(\bar{\Psi}\) later. Since we will only find no non-vanishing contributions for the \(c^\dagger\) at this order, we can focus on

\[
\bar{\Psi}(x_1) \Psi(x_1) \bar{\Psi}(x_2) \Psi(x_2) : b^\dagger_{p_1} b^\dagger_{p_2} |0\rangle
\]

\[
= - \frac{1}{(2\pi)^6} \int \frac{d^3 k_1 d^3 k_2}{E_{k_1} E_{k_2}} \overline{\Psi}(x_1) \cdot u^a(k_1) (\Psi(x_2) \cdot u^a(k_2)) e^{ik_1 \cdot x_1 - ik_2 \cdot x_2} \sqrt{4E_{k_1} E_{k_2}} b^\dagger_{k_1} b^\dagger_{k_2} b^\dagger_{p_1} b^\dagger_{p_2} |0\rangle
\]

where the \(-\) sign is from commuting \(\Psi\) past the \(\bar{\Psi}\). Let us now anticommutate the \(b^\dagger\)s past the \(b^\dagger\)s \(^{18}\) to obtain

\[
= - \frac{1}{\sqrt{4 E_{p_1} E_{p_2}}} \left[ \overline{\Psi}(x_1) \cdot u^a(p_2) \langle \Psi(x_2) \cdot u^a(p_1)e^{-i p_2 \cdot x_1 - i p_1 \cdot x_2} - \langle \Psi(x_1) \cdot u^a(p_1) \rangle \overline{\Psi}(x_2) \cdot u^a(p_2) e^{-i p_1 \cdot x_2 - i p_2 \cdot x_1} |0\rangle \right]
\]

Note again the relative minus sign between the two terms. What happens when we contrast this with \(|f\rangle\)? Let’s see the first term

\[
\langle 0 \vert \hat{S}_{q_2} b^\dagger_{p_1} \hat{S}_{q_1} \langle \Psi(x_1) \cdot u^a(p_2) \rangle \overline{\Psi}(x_2) \cdot u^a(p_1) |0\rangle = \frac{e^{i q_2 \cdot x_1 + i q_1 \cdot x_2}}{2 \sqrt{E_{q_1} E_{q_2}}} \langle \hat{u}^a(q_1) \cdot u^a(p_2) \rangle \langle \hat{u}^a(q_2) \cdot u^a(p_1) \rangle
\]

the second term contributes the same and cancels the 1/2 in front of Eq. (7.81). The factors of \(1/\sqrt{E}\) cancel the relativistic state normalizing. Combining everything, we obtain

\[
- (-i \lambda)^2 \int \frac{d^4 x_1 d^4 x_2 d^4 k}{(2\pi)^4} \frac{ie^{i (x_1 - x_2)}}{k^2 - \mu^2 + i\varepsilon} \left[ \langle \hat{u}^a(q_1) \cdot u^a(p_2) \rangle \langle \hat{u}^a(q_2) \cdot u^a(p_1) \rangle e^{i (q_2 - p_1) \cdot x_1 + i (q_1 - p_2) \cdot x_2} - \langle \hat{u}^a(q_2) \cdot u^a(p_2) \rangle \langle \hat{u}^a(q_1) \cdot u^a(p_1) \rangle e^{i (q_1 - p_1) \cdot x_1 + i (q_2 - p_2) \cdot x_2} \right]
\]

where have used the \(\phi\)-propagator. Integrating over \(x_1\) and \(x_2\), we obtain

\[
- \int \frac{d^4 k}{k^2 - \mu^2 + i\varepsilon} \left[ \langle \hat{u}^a(q_1) \cdot u^a(p_2) \rangle \langle \hat{u}^a(q_2) \cdot u^a(p_1) \rangle \delta^{(4)}(p_1 - q_1 + k) \delta^{(4)}(p_2 - q_2 + k) - \langle \hat{u}^a(q_2) \cdot u^a(p_2) \rangle \langle \hat{u}^a(q_1) \cdot u^a(p_1) \rangle \delta^{(4)}(p_1 - q_2 + k) \delta^{(4)}(p_2 - q_1 + k) \right]
\]

Almost done. We use the usual normalization

\[
\langle f \vert S - 1 \vert i \rangle = i A (2\pi)^4 \delta^{(4)}(p_1 + p_2 - q_1 - q_2)
\]

we find

\[
i A = (-i \lambda)^2 \left[ \frac{\langle \hat{u}^a(q_2) \cdot u^a(p_2) \rangle \langle \hat{u}^a(q_1) \cdot u^a(p_1) \rangle}{(p_2 - q_1)^2 - \mu^2 + i\varepsilon} - \frac{\langle \hat{u}^a(q_1) \cdot u^a(p_2) \rangle \langle \hat{u}^a(q_2) \cdot u^a(p_1) \rangle}{(p_1 - q_1)^2 - \mu^2 + i\varepsilon} \right]
\]

which is the final result.
7.8 Feynman rules for Dirac Spinors

The calculation above is really cumbersome and you will hopefully never have to redo it this way ever again. Luckily introducing Feynman rules simplifies the calculation significantly and we will see that we will be able to derive the final result in a couple of easy steps.

1. Propagators:

\[ \phi(x)\phi(y) = \frac{i}{p^2 - \mu^2 + i\varepsilon} = \hspace{2cm} \]
\[ \Psi(x)\Psi(y) = \frac{i(p + m)}{p^2 - m^2 + i\varepsilon} = \hspace{2cm} \]

2. Vertices:

\[ -i\lambda = \]

3. External legs incoming

\[ p = u^*(p) \]
\[ p = \bar{v}^*(p) \]

4. External legs outgoing

\[ p = \bar{u}^*(p) \]
\[ p = v^*(p) \]

5. Impose momentum conservation at each vertex.

6. Integrate over each undetermined momentum \( k \) with \( \int \frac{d^4k}{(2\pi)^4} \).

7. Figure out the overall sign of the diagram.

8. A closed fermion loop always gives a factor of \(-1\) and the trace of a product of Dirac matrices.\(^{19}\)

We draw arrows on fermion lines to represent the direction of particle-number flow. The momentum assigned to a fermion propagator flows in the direction of this arrow. For external antiparticles, however, the momentum flows opposite to the arrow. We show this explicitly by drawing a second arrow next to the line.

\(^{19}\) We will see an example of this below.
7.8.1 Example: Nucleon scattering $\Psi \Psi \rightarrow \Psi \Psi$

The two diagrams are

\[
\begin{align*}
\Psi & \rightarrow \Psi \\
\uparrow & \quad \downarrow \\
p & \quad q \\
\quad & \quad + \\
\quad & \quad \uparrow \downarrow \\
p' & \quad q'
\end{align*}
\]

The amplitude for the diagram is

\[
iA = (-ig)^2 \left[ \bar{u}(q') \cdot u(q) \cdot \frac{i}{(q-q')^2 - \mu^2} \bar{u}(p') \cdot u(p) \
- \bar{u}(p') \cdot u(q) \cdot \frac{i}{(p' - q)^2 - \mu^2} \bar{u}(q') \cdot u(p) \right]
\]

(7.83)

Where we have suppressed the spinor spin labels. The minus sign between the diagrams is a reflection of Fermi-Dirac statistics.

7.8.2 Example: Nucleon-Anti-Nucleon scattering $\Psi \bar{\Psi} \rightarrow \Psi \bar{\Psi}$

The two diagrams are

\[
\begin{align*}
\Psi & \rightarrow \Psi \\
\uparrow & \quad \downarrow \\
p & \quad q \\
\quad & \quad + \\
\quad & \quad \uparrow \downarrow \\
p' & \quad q'
\end{align*}
\]

with the Feynman rules we have

\[
iA = (-i\lambda)^2 \left[ \bar{\psi}(q') \cdot u(p) \cdot v(q) \cdot \bar{u}(p') \cdot \bar{v}(q') \
+ \bar{v}(q) \cdot u(p) \cdot \bar{u}(p') \cdot v(q') \right]
\]

(7.84)

There is a relative minus sign between the two diagrams, which is a bit subtle to figure out. You can convince yourself that it is correct, by repeating the relevant steps using the explicit Wick technology above.

7.8.3 Example: Meson Scattering

Finally, in $\phi \phi \rightarrow \phi \phi$ scattering, the leading contribution appears first at 1-loop. The amplitude for the diagram is
\[ iA = -(-i\lambda)^4 \int \frac{d^4k}{(2\pi)^4} \text{Tr} \left[ \frac{k + m}{k^2 - m^2} \frac{k + p + m}{(k + p)^2 - m^2} \frac{k' - p + m}{k' + m + q'} \right] \]

where we have suppressed the \( + i\varepsilon \) piece in the denominator. Note, that the integral diverges logarithmically\(^{20}\) In the \( k \to \infty \) limit, the integral behaves like

\[ iA \sim \int \frac{d^4k}{k^4} \sim \ln \Lambda \]

You will have to come to the QFT course next semester to learn how to make sense of results like these. There is an overall minus sign sitting in front the of the amplitude.

In complicated diagrams, we can often find the minus signs, by noting that the product \( \langle \bar{\Psi} \Psi \rangle \) or any other pair of fermions is a commuting object. So,

\[ \ldots \langle \bar{\Psi} \Psi \rangle_x \langle \bar{\Psi} \Psi \rangle_y \langle \bar{\Psi} \Psi \rangle_z \ldots = \ldots (+1) \langle \bar{\Psi} \Psi \rangle_x \langle \bar{\Psi} \Psi \rangle_z \langle \bar{\Psi} \Psi \rangle_y \ldots \]

\[ \ldots S_F(x - z) S_F(z - y) \ldots \]

In a \textbf{closed loop} of \( n \) fermion propagators we have

\[ - - \]

which involves the contraction

\[ \bar{\Psi}_\alpha(x) \Psi_\alpha(x) \bar{\Psi}_\beta(y) \Psi_\beta(y) = - \bar{\Psi}_\beta(y) \bar{\Psi}_\alpha(x) \Psi_\alpha(x) \Psi_\beta(y) \]

\[ = - \text{Tr}[S_F(y - x) S_F(x - y)] \]

This is a general result: a closed fermion loop results in a factor of \((-1)\) and a \textbf{trace} of a string of \( \gamma \) matrices.

\subsection*{7.8.4 Example: the Yukawa potential}

We saw in Sec. ?? that the exchange of a real scalar between to charged scalars gives rise to always attractive Yukawa potential between two spin zero particles. Do we still find the same for the scalar force between two spin 1/2 particles?

We will again take the \textbf{non-relativistic} limit \( p \ll m \) which implies for the spinors

\[ u^s(p) = \begin{pmatrix} \sqrt{p \cdot \sigma} \xi^s \\ -\sqrt{p \cdot \sigma} \xi^s \end{pmatrix} \rightarrow \sqrt{m} \begin{pmatrix} \xi^s \\ -\xi^s \end{pmatrix} \]

and

\[ v^s(p) = \begin{pmatrix} \sqrt{p \cdot \sigma} \xi^s \\ -\sqrt{p \cdot \sigma} \xi^s \end{pmatrix} \rightarrow \sqrt{m} \begin{pmatrix} \xi^s \\ -\xi^s \end{pmatrix} \]

If the two interacting particles are distinguishable only the first one in Eq. (7.82) contributes. In this limit, the spinor contractions in the \( \psi\psi \rightarrow \bar{\psi}\bar{\psi} \) scattering are

\[ u^s(q') \cdot u^s(q) = 2m \delta ss' \]
and the amplitude becomes

\[ i\mathcal{A} = -i(-i\lambda)^2(2m) \left( \frac{\delta_{s's'}\delta_{r'r}}{(p - p')^2 + \mu^2} \right) \]

We see that the spin is conserved at each vertex in the non-relativistic limit (due to the \( \delta \) factors) and that the momentum dependence is exactly the same as in the bosonic case Eq. (6.47), telling us that the particles feel an **attractive Yukawa potential**,

\[ U(r) = -\frac{\lambda^2 e^{-\mu r}}{4\pi r} \]

If we repeat the calculation for \( \Psi \bar{\Psi} \rightarrow \bar{\Psi} \Psi \), two minus signs cancel and the Yukawa interaction once again leads to an attractive potential.
8

Quantum Electro Dynamics

8.1 The Feynman Rules for Quantum Electrodynamics

We replace the scalar particle $\phi$ with a vector particle $A_{\mu}$ and replace the Yukawa interaction Hamiltonian with

$$H_{\text{int}} = \int d^3x \bar{\Psi} \gamma^\mu \Psi A_{\mu}$$

What are the Feynman rules? We can easily guess, even though the theory behind them is quite involved.

1. Interaction vertex:

$$-ie\gamma^\mu$$

2. Propagators:

$$A_{\mu}(x)A_{\nu}(y) = \frac{-i\eta_{\mu\nu}}{q^2 + i\varepsilon}$$

3. External photon lines:

   Outgoing

$$p = \epsilon_{\mu}(p)$$

   Incoming

$$p = \epsilon_{\mu}(p)$$

We draw photons as wavy lines. The symbol $\epsilon_{\mu}(p)$ stands for the polarization vectors of the initial or final state photons. The polarization vectors are always of the form

$$\epsilon^\mu = \begin{pmatrix} 0 \\ \epsilon \end{pmatrix}, \quad \text{with} \quad p \cdot \epsilon = 0 \quad (8.1)$$
If \( p \) is along the z-axis, the right- and left-handed polarization vectors are

\[
\epsilon^\mu = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ \pm i \\ 0 \end{pmatrix}
\]

\( 8.2 \) Maxwell’s equations

The Maxwell-Lagrangian in the absence of any sources is simply

\[
\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}
\]

with the field strength

\[
F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu
\]

and the equations of motions are

\[
\partial_\mu \left( \frac{\partial \mathcal{L}}{\partial (\partial_\mu A_\nu)} \right) = -\partial_\mu F^{\mu\nu} = 0
\]

Additionally, the field strength satisfies the **Bianchi** identities

\[
\partial_\lambda F^{\mu\nu} + \partial_\mu F^{\nu\lambda} + \partial_\nu F^{\lambda\mu} = 0
\]

The version of Maxwell’s equation you learn in pre-school is recovered atop with the fields

\[
A^\mu = \begin{pmatrix} \phi \\ A \end{pmatrix}
\]

with the electric and magnetic fields \( E \) and \( B \)

\[
E = -\nabla \phi - \frac{\partial A}{\partial t}
\]

and

\[
B = \nabla \times A
\]

which in terms of the field strength is

\[
F_{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix}
\]

and the Bianchi identity results in two Maxwell equations which are unchanged even in the presence of sources

\[
\nabla \cdot B = 0
\]

\[
\frac{\partial B}{\partial t} = -\nabla \times E
\]

The equations of motions give the two remaining Maxwell equations

\[
\nabla \cdot E = 0
\]

\[
\frac{\partial E}{\partial t} = \nabla \times B
\]

Note, the interesting electro-magnetic duality \( E \rightarrow B \) and \( B \rightarrow -E! \)
8.3 Gauge Symmetry

We have used $A_\mu$ to describe a massless vector field, which would seem to imply that since the gauge field has 4 components, we naively count 4 degrees of freedom. Yet we know has only two degrees of freedom, which are the polarization states we’ve used in the Feynman rules. How can this be resolved? There are two mechanisms at work which ensure this.

- The field $A_0$ is not dynamical: It has no kinetic term $A_0$ in the Lagrangian. If we fix $A_i$ and $\dot{A}_i$ at initial some time $t_0$, then $A_0$ is fully determined by the equation of motion $\nabla \cdot E = 0$ which we can expand out to obtain

$$\nabla^2 A_0 + \nabla \cdot \frac{\partial A}{\partial t} = 0$$

which has the solution

$$A_0(x) = \int d^3x' \frac{\nabla \cdot \frac{\partial A(x')}{\partial x}}{4\pi |x - x'|} \tag{8.4}$$

which you can convince yourself is correct using the arguments around Eq. (4.95) and the Green’s function method. So $A_0$ is clearly dependent and we do not get to specify it fixing the initial conditions. This reduces our the degrees of freedom from 4 to 3.

- The Maxwell Lagrangian has a very large redundancy,

$$A_\mu(x) \rightarrow A_\mu(x) + \partial_\mu \alpha(x) \tag{8.5}$$

for any function $\alpha(x)$. We will require that $\alpha(x)$ vanishes fast enough at $x \rightarrow \infty$. This is a gauge symmetry. The field strength is invariant under it

$$F_{\mu\nu} \rightarrow \partial_\mu (A_\nu + \partial_\nu \alpha) - \partial_\nu (A_\mu + \partial_\mu \alpha) = F_{\mu\nu}$$

We seem to have a theory with an infinite number of symmetries, each for each local choice of function $\alpha(x)$. All our other internal symmetries where global and acted the same at every point in spacetime, e.g.

$$\psi \rightarrow e^{i\beta} \psi$$

with $\beta = \text{const.}$. These symmetries lead to selection rules or conservation laws, thanks to Noether’s theorem. Do we now have infinitely many conserved currents?

No! A symmetry takes a physical state to different physical state with the same properties, the gauge symmetry however is a redundancy in our description. The two states connected by a gauge symmetry have to be identified, they are the same physical state. You can see this also in the Maxwell’s equation which are not sufficient to specify the time evolution of $A_\mu$

$$(\eta_{\mu\nu}(\partial^\lambda \partial_\lambda) - \partial_\mu \partial_\nu)A^\nu = 0$$

E.g. from counting states and calculating the thermodynamics of a blackbody in thermodynamics.
Now, the differential operator $(\eta_{\mu\nu}(\partial^\lambda \partial_\lambda) - \partial_\mu \partial_\nu)$ is \textbf{not invertible}, since it has a zero eigenvector for functions of the form $\partial_\mu \lambda$.

Fixing initial data, we can not uniquely determine $A_\mu$ since there no way to distinguish between

$$A_\mu, \quad \text{and} \quad A_\mu + \partial_\mu \lambda$$

If we are willing to identify $A_\mu$ and $A_\mu + \partial_\mu \lambda$ as the \textbf{same physical state}, then this is not a problem anymore.

How about formulating the theory solely in $E$ and $B$? Lorentz-invariance would not be manifest, but we might be able to deal with this. However, this is not advisable given the relevance of $A_\mu$ for quantum mechanical effects like the \textbf{Aharonov-Bohm effect}.

Further, we would like to describe charged gauge fields, which also requires generalizations of $A_\mu$.

We discover an enlarged phase space, foliated by gauge orbits. To remove the redundancy, we need to pick one representative from each equivalent gauge orbit. We call these different configuration of a physical state, different \textbf{gauges}. We can pick gauges as we please, but as for coordinate systems, there are better and less suited choices depending on the problem.

### 8.3.1 Gauges

We start with the \textbf{Lorentz gauge}\(^1\)

$$\partial_\mu A^\mu = 0$$

This does not yet pick a unique representation from the gauge orbit, since we are free to make further gauge transformations which are functions that satisfy

$$\partial_\mu \partial^\mu \lambda(x) = 0$$

which certainly has non-trivial solutions. The advantage of the Lorentz gauge condition is clear from its name: it is manifestly Lorentz invariant.

Another useful gauge, is the \textbf{Coulomb gauge}

$$\nabla \cdot A = 0$$

there is still options to choose harmonic functions which satisfy

$\nabla^2 \lambda = 0$. Since Eq. (8.4) fixes $A^0$, we find that now\(^2\)

$$A_0 = 0$$

Coulomb gauge is not manifestly Lorentz invariant, but it clearly shows the degrees of freedom of the physical system: The 3 degrees of freedom satisfy a single constraint $\nabla \cdot A = 0$, leaving us with just 2 independent degrees of freedom.

---

\(^1\) Other popular choices are $A^0 = 0$ (temporal gauge) and $A^3 = 0$ (axial gauge)

\(^2\) This will not hold in the presence of sources, as you know from electrostatics, since in this case, e.g. a charge induces a potential for $A_0 = \phi$
8.4 Quantizing the Electro-magnetic field

We will now try to quantize the electric field in Coulomb and in Lorentz gauge. A common subtlety to both methods is due to the absence of a canonical momentum for $A_0$:

\[ \pi^0 = \frac{\partial L}{\partial (\partial_t A_0)} = 0 \quad (8.8) \]
\[ \pi^i = \frac{\partial L}{\partial (\partial_t A_i)} = -F^{0i} = E^i \quad (8.9) \]

We again see that $A_0$ is not a dynamical field. The Hamiltonian is

\[ H = \int d^3x (\pi^i A_i - L) = \int d^3x \left( \frac{1}{2} E \cdot E + \frac{1}{2} B \cdot B - A_0 (\nabla \cdot E) \right) \]

We see that $A_0$ acts as a Lagrange multiplier, which has the equation of motion

\[ \nabla \cdot E = 0 \]

which is just enforcing Gauss’ law. What do the different gauge fixing conditions mean for the system?

8.4.1 Coulomb gauge

In Coulomb gauge, the equation of motion for $A$ is

\[ \partial_\mu \partial^\mu A = 0 \]

which we solve as usual by a Fourier-ansatz

\[ A(x) = \int \frac{d^3p}{(2\pi)^3} \xi(p) e^{ip \cdot x} \]

with $p^2 = 0$ or $p_0^2 = |p|^2$. The Coulomb gauge condition $\nabla \cdot A = 0$ requires

\[ \xi(p) \cdot p = 0 \]

which means that $\xi(p)$ is orthogonal to the direction of motion of the gauge field $p$. We can choose it to be a linear combination of two ortho-normal vectors $\epsilon_r$ with $r = 1, 2$ which satisfy

\[ \epsilon_r(p) \cdot p = 0 \]

and

\[ \epsilon_r(p) \cdot \epsilon_s(p) = \delta_{rs} \]

These are the two polarization states of the photon. Now we attempt to quantize and we turn the classical Poisson brackets into commutator for operator valued fields. Naively (and wrongly, we try for the equal-time commutator

\[ [A_i(x), E^j(y)] = i\delta^j_i \delta^{(3)}(x - y) \quad \text{(wrong)} \]
and

$$[A_i(x), A_j(y)] = [E^i(x), E^j(y)] = 0$$

This is not correct! Why? It does not work with the gauge and Gauss constraints. We need to have imposed on the operators

$$\nabla \cdot A = \nabla \cdot E = 0$$

but if we evaluate the operator relations above, we find

$$[\nabla_x \cdot A(x), \nabla_y \cdot E(y)] = i \nabla^2 \delta^i_j (x - y) \neq 0$$

You can see this problem already in the classical Poisson bracket structure.³ The correct Poisson bracket structure implies different commutation relations in Coulomb gauge

$$[A_i(x), E^j(y)] = i \left( \delta^i_j - \frac{\partial_i \partial_j}{\nabla^2} \right) \delta^{(3)}(x - y) \quad (8.10)$$

Let us check if this is consistent with the constraints. We write the $\delta$ function in momentum space

$$[A_i(x), E^j(y)] = i \int \frac{d^3 p}{(2\pi)^3} \left( \delta^i_j - \frac{p_i p_j}{p^2} \right) e^{i p \cdot (x - y)}$$

Applying e.g. the Coulomb gauge fixing

$$[\partial_x^i A_i(x), E^j(y)] = i \int \frac{d^3 p}{(2\pi)^3} i p_i \left( \delta^i_j - \frac{p_i p_j}{p^2} \right) e^{i p \cdot (x - y)}$$

$$= i \int \frac{d^3 p}{(2\pi)^3} i p_j \left( \frac{p^2 p_j}{p^2} \right) e^{i p \cdot (x - y)} = 0$$

We expand in modes again as usual

$$A(x) = \sum_{r=1}^2 \int \frac{d^3 p}{(2\pi)^3} \frac{1}{\sqrt{2|p|}} \epsilon_r(p) \left[ a^r_p e^{i p \cdot x} + a^{r\dagger}_p e^{-i p \cdot x} \right]$$

$$E(x) = \sum_{r=1}^2 \int \frac{d^3 p}{(2\pi)^3} \frac{(-i) \sqrt{|p|}}{\sqrt{2}} \epsilon_r(p) \left[ a^r_p e^{i p \cdot x} - a^{r\dagger}_p e^{-i p \cdot x} \right]$$

with the before define polarization vectors.

$$\epsilon_r(p) \cdot p = 0 \quad \text{and} \quad \epsilon_r(p) \cdot \epsilon_s(p) = \delta_{rs}$$

You should convince yourself that the commutation relations imply for the ladder operators

$$[a^r_p, a^s_{q\dagger}] = (2\pi)^3 \delta^{rs} \delta^{(3)}(p - q)$$

and

$$[a^{r\dagger}_p, a^s_q] = [a^{r\dagger}_p, a^s_{q\dagger}] = 0$$

³ See e.g. P. Ramond - QFT, or Dirac "Lectures on Quantum mechanics".
where you will have to use the outer product (or completeness relation) for the polarization vectors

\[ \sum_{i=1}^{2} \epsilon_i^*(p) \epsilon_i(p) = \delta^{ij} - \frac{p_i p_j}{p^2} \]

We can confirm this equation by acting on both sides with a complete basis \((\epsilon_1(p), \epsilon_2(p), p)\).

We substitute the expansion into the Hamiltonian and obtain after normal ordering

\[ H = \int \frac{d^3p}{(2\pi)^3} |p| \sum_{r=1}^{2} a_{r}^\dagger a_{r} \]

Quantizing in the Coulomb gauge gave us direct insight into the physical degrees of freedom. Lorentz invariance is however not manifest anymore. We can see this looking at the propagator in the Heisenberg picture

\[ D_{ij}^{tr}(x-y) \equiv \langle 0| T A_i(x) A_j(y) |0 \rangle = \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 + i\epsilon} \left( \delta^{ij} - \frac{p_i p_j}{p^2} \right) e^{-ix\cdot p} \]

Where explicitly denote with \(tr\), the transverse part of the photon. Let us therefore explore the manifestly invariant Lorentz gauge.

### 8.4.2 Lorentz Gauge

The Lorentz gauge condition is

\[ \partial_{\mu} A^\mu = 0 \]

which implies for the equations of motion

\[ \partial^\mu \partial_{\mu} A^\nu = 0 \]

We will depart from the procedure enforcing Coulomb gauge\(^4\) and will modify the Lagrangian such that the equations of motion are directly in Lorentz gauge. With

\[ \mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2} (\partial_{\mu} A^\mu)^2 \]

we obtain

\[ \partial_{\mu} F^{\mu\nu} + \partial^\nu (\partial_{\mu} A^\mu) = \partial^\nu \partial_{\mu} A^\nu = 0 \]

We can be more general and choose

\[ \mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2\alpha} (\partial_{\mu} A^\mu)^2 \]

with arbitrary and constant \(\alpha\). We confusingly call \(\alpha = 1\) the **Feynman gauge** and \(\alpha = 0\) the Landau gauge.

We will now firstly quantize the theory and only secondly enforce the gauge constraint. Thankfully, both conjugate momenta are

\(^4\) This can be motivated much more cleanly using the path integral formalism and a trick by Fadeev and Popov.
non-vanishing
\[ \pi^0 = \frac{\partial L}{\partial (\partial_0 A_0)} = -\partial_\mu A^\mu \] (8.11)
\[ \pi^i = \frac{\partial L}{\partial (\partial_i A_i)} = \partial^A A^0 - \partial^0 A^i \] (8.12)

We impose the equal-time commutation relations in Lorentz-gauge
\[ [A_\mu(x), \pi_\nu(y)] = i \eta_{\mu\nu} \delta^{(3)}(x - y) \] (8.13)

and
\[ [A_\mu(x), A_\nu(y)] = [\pi_\mu(x), \pi_\nu(y)] = 0 \]

Expanding again
\[ A_\mu(x) = \sum_{\lambda=0}^3 \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{|p|}} \epsilon^\lambda(p) [\alpha^\lambda p e^{ip \cdot x} + \alpha^\lambda + e^{-ip \cdot x}] \]
\[ \pi_\mu(x) = \sum_{\lambda=0}^3 \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{|p|}} \epsilon^\lambda(p) [\alpha^\lambda p e^{ip \cdot x} - \alpha^\lambda + e^{-ip \cdot x}] \]

Note that we here have \((+ i)\) in the momentum, rather than the usual \((- i)\). The reason is that the conjugate momentum is
\[ \pi^\mu = -\dot{A}^\mu + \ldots \]

We now have to deal with four polarization vectors! This is much worse than the two physical polarizations in Coloumb gauge. We choose \(\epsilon^0\) time-like and \(\epsilon^i\) to be space-like. We choose the normalization
\[ \epsilon^\lambda \cdot \epsilon^{\lambda'} = \eta^{\lambda\lambda'} \]
which implies
\[ (\epsilon_\mu)^\lambda (\epsilon_\nu)^{\lambda'} \eta_{\lambda\lambda'} = \eta_{\mu\nu} \]

We choose two of the space-like polarizations to be transverse to the momentum \(p^\mu = (|p|, p)\)
\[ \epsilon^1 \cdot p = \epsilon^2 \cdot p = 0 \] (8.14)

The third vector is a longitudinal polarization. If \(p_{(z)}^\mu \propto (1, 0, 0, 1)\), we would have
\[ \epsilon^0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \epsilon^1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \epsilon^2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \epsilon^3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \]

We define for a general four-momentum \(p^\mu \propto (|p|, p)\),
\[ \epsilon^0(p) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \epsilon^1(p) = \frac{1}{|p|} \begin{pmatrix} 0 \\ p^1 \\ p^2 \\ p^3 \end{pmatrix}, \quad \epsilon^r(p) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \epsilon^r \end{pmatrix} \]
with \(\epsilon^r(p) \cdot p = \epsilon^r(p) \cdot p = 0, \quad r = 1, 2\)
For later use, we note for the time-like and longitudinal polarizations that
\[ \epsilon^0(p) \cdot p = |p| \quad \epsilon^3(p) \cdot p = -|p| \] (8.15)

We derive the ladder operator commutation relations using the field expansion to obtain
\[ [a^\lambda_p, a^\lambda'_{q\dagger}] = -\eta^{\lambda\lambda'} (2\pi)^3 \delta^{(3)}(p-q) \]

with all other commutators vanishing as usual. The minus sign is problematic! For space-like \( \lambda, \lambda' = 1, 2, 3 \) we are in well-known territory
\[ [a^\lambda_p, a^\lambda_{q\dagger}] = (2\pi)^3 \delta^{\lambda\lambda'} (p-q), \quad \lambda, \lambda' = 1, 2, 3 \]

but for time-like ladder operators, we have
\[ [a^0_p, a^0_{q\dagger}] = -(2\pi)^3 \delta^{(3)}(p-q) \]

This means trouble. Let us see what it implies e.g. for one-particle states. With the vacuum
\[ a^\lambda_p |0\rangle = 0 \]

we can define (yet to be normalized) one-particle states as usual
\[ |p, \lambda\rangle = a^\lambda_{p\dagger} |0\rangle \]

For space-like \( \lambda = 1, 2, 3 \) we are in the green, but for the time-like polarization \( \lambda = 0 \), the state \( |p, \lambda = 0\rangle \) has negative norm
\[ \langle p, 0 | q, 0 \rangle = \langle 0 | a^0_p a^0_{q\dagger} |0\rangle = -(2\pi)^3 \delta^{(3)}(p-q) \]

This strange. Negative norm would imply negative probabilities! We put ourselves in danger already from the start as we can see from the kinetic term for \( A_0 \) which has the wrong sign
\[ L = \frac{1}{2} \dot{A}^2 - \frac{1}{2} \dot{A}_0^2 + \ldots \]

How can we make sense of this?

We have yet to impose the gauge constraint \( \partial_\mu A^\mu = 0 \) on the theory. It will be responsible to remove the space-like negative norm states and remove all but two polarizations. How can we implement the constraint in a quantum theory?

- Let us start with the strongest possible condition. We impose \( \partial_\mu A^\mu = 0 \) as an operator equation. This does not work, because it is not compatible with the commutation relations Eq. (8.13), since the conjugate momentum Eq. (8.11)
\[ \pi^0 = \frac{\partial L}{\partial (\dot{A}_0)} = -\partial_\mu A^\mu \]

would vanish universally.
A weaker option would be to impose it on a subset of the **Hilbert space** instead of on the operators. We could imagine splitting the Hilbert space into ”good” states $|\psi\rangle$ and a ”bad” states which somehow are not part of physical observables. How can we define ”good” states. We could impose

$$\partial_\mu A_\mu|\psi\rangle = 0$$  \hspace{1cm} (8.16)

on all the ”good” states. However this does not work either. It is still too strong, since we can e.g. decompose $A_\mu = A_\mu^+ + A_\mu^-$ with

$$A_\mu^+(x) = \sum_{\lambda=0}^3 \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2|p|}} \epsilon^\lambda(p) a^\lambda_p e^{-ip \cdot x}$$

$$A_\mu^-(x) = \sum_{\lambda=0}^3 \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2|p|}} \epsilon^\lambda(p) a^{\dagger \lambda}_p e^{ip \cdot x}$$

The vacuum satisfies

$$A_\mu^+|0\rangle = 0$$

by definition. But we get

$$\partial^\mu A_\mu^-|0\rangle \neq 0$$

which shows that the vacuum as a physical state isn’t even one of the good states.

The solution: We want the vacuum as one of the good physical states. We therefore ask physical states $|\psi\rangle$ to only satisfy the **Gupta-Bleuler** condition

$$\partial^\mu A_\mu^+|\psi\rangle = 0$$ \hspace{1cm} (8.17)

and the adjoint equation

$$\langle \psi|\partial^\mu A_\mu^- = 0$$

With this we make sure that

$$\langle \psi'|\partial^\mu A_\mu^- + \partial^\mu A_\mu^+|\psi\rangle = \langle \psi'|\partial^\mu A_\mu^-|\psi\rangle = 0$$

holds. Since the constraint is linear, the physical states span the physical Hilbert space $H_{\text{phys}}$. How does the physical Hilbert space look like?

We consider a basis for the Fock space. We decompose any element of this basis into

$$|\psi\rangle = |\varphi_T\rangle |\varphi_{L-0}\rangle$$

where $|\varphi_T\rangle$ contains the transverse photons (created by $a_1^\dagger_p$ and $a_2^\dagger_p$) and $|\varphi_{L-0}\rangle$ contains time-like and longitudinal photons (created by $a_0^\dagger_p$ and $a_3^\dagger_p$).

---

5 The problem with the Gupta-Bleuer condition is that it is not suited for non-abelian gauge theories due to the non-linearity of the free field equations (they contain terms $\sim A_1^3, A_4^4$). The modern approach replacing this QED-only ‘hack’ is called **BRST** symmetry (Becchi, Rouet, Stora and Tyutin) and will be developed in the QFT lectures. One identifies a nilpotent symmetry generator which commutes with the Hamiltonian $H$ which allows to split the eigenstates of $H$ into unphysical and physical sub-sectors which have zero inner product with each other.
The Gupta-Bleuer condition Eq. (8.17) therefore reads with the polarization condition for the transverse states Eq. (8.14) and Eq. (8.15) as

$$(a_p^0 - a_p^3) |\varphi_{L=0}⟩ = 0 \quad (8.18)$$

What does this mean? Condition Eq. (8.18) requires all physical states to contain pairs of timelike and longitudinal photons: for each timelike photon with $p$, it must also contain a longitudinal photon with the same $p$. We can write $|\varphi_{L=0}⟩$ as a sum over states $|\varphi^{(n)}⟩$ containing $n$ pairs of timelike and longitudinal photons

$$|\varphi_{L=0}⟩ = \sum_{n=0}^{∞} c_n |\varphi^{(n)}⟩ \quad (8.19)$$

where $|\varphi^{(0)}⟩ = |0⟩$ is just the vacuum.

One can show that Eq. (8.18) indeed decouples the negative norm states and the all the remaining timelike and longitudinal photons have zero norm

$$\langle \varphi^{(n)} | \varphi^{(m)} \rangle = δ_{n,0} δ_{m,0}$$

We have therefore defined an inner product on the physical Hilbert space which is positive semi-definite. This is clearly a step in the right direction. We will now have to find a way to deal with the zero-norm states.

We understand the zero-norm states if we treat them as gauge-equivalent to the vacuum. If two states that differ only by pairs of timelike and longitudinal photons, $|\varphi^{(n)}⟩$ with $n \geq 1$, they describe the same physical state.

We need to check then that they give the same result for all physical observables. Let us start with the Hamiltonian, which you can easily compute

$$H = \int \frac{d^3p}{(2\pi)^3} |p| \left[ \sum_{i=1}^{3} a_p^{i+} a_p^i - a_p^0 a_p^0 \right]$$

The Gupta-Bleuer condition Eq. (8.18) guarantees that

$$\langle \psi | a_p^{3+} a_p^3 | \psi⟩ = \langle \psi | a_p^0 a_p^0 | \psi⟩$$

and the contributions of of timelike and longitudinal photons cancel each other in the Hamiltonian. Further, the Hamiltonian evaluated on physical states satisfying the Gupta-Bleuer condition is positive definite and contains only the contributions from transverse photons, as it should be.

One can show that expectation values of gauge-invariant operators evaluated on physical states are independent of the coefficients $c_n$ of Eq. (8.19).
8.4.3 The propagator

The propagator for the photon in Lorentz gauge is given by

\[ \langle 0 | T A_\mu(x) A_\nu(y) | 0 \rangle = \int \frac{d^4p}{(2\pi)^4} \frac{-i\eta_{\mu\nu}}{q^2 + i\varepsilon} e^{-ip(x-y)} \]

Alternatively, you can show that for general gauges with arbitrary \( \alpha \) in Eq. (??), we would obtain in momentum space

\[ -i \frac{q^2}{q^2 + i\varepsilon} \left( \eta_{\mu\nu} + (\alpha - 1) \frac{p_\mu p_\nu}{p^2} \right) \]

8.5 Coupling gauge fields to matter

We now want to couple the gauge field \( A_\mu \) to other fields, like scalars or fermions (or more poetically, coupling light to matter). Let us see how we can do this. We know that the simplest coupling has to have the form

\[ L = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - A_\mu R^\mu \]  

(8.20)

What are the conditions on \( R^\mu \)? We know that we need the gauge redundancy to find the correct number of degrees of freedom. The equations of motions are

\[ \partial_\nu F^{\mu\nu} = R^\mu \]

Let us apply \( \partial_\nu \) on both sides

\[ \partial_\nu \partial_\mu F^{\mu\nu} = 0 = \partial_\nu R^\mu \]

In other words \( R^\mu \) must be a conserved current! We write \( R^\mu = j^\mu \) and recognize this as the usual conservation equation

\[ \partial_\mu j^\mu = 0 \]

We have already encountered a large number of these conserved currents \( j^\mu \). Which are the Noether currents related to the gauge symmetry for the photon field?

8.5.1 Coupling photons to Dirac fermions

Let us try with the \( U(1) \) global, internal symmetry of the Dirac Lagrangian

\[ \Psi \rightarrow e^{i\alpha} \Psi \]

which with Eq. (7.60) gives rise to a current

\[ j^\mu = \overline{\Psi} \gamma^\mu \Psi \]  

(8.21)

Now let us couple this to the gauge field \( L \supset A_\mu j^\mu \) to obtain together with the kinetic term of the Dirac spinor

\[ L = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \overline{\Psi} \phi \Psi - e \overline{\Psi} \gamma^\mu \Psi A_\mu \]  

(8.22)
where I have introduced a yet to be specified constant \( e \). We can rewrite this as

\[
\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + i \overline{\Psi} \slashed{D} \Psi \tag{8.23}
\]

with the covariant derivative

\[
D_\mu = \partial_\mu + ieA_\mu \tag{8.24}
\]

Under a gauge transformation \( A_\mu(x) \to A_\mu(x) + \partial_\mu \lambda(x) \) we find that the Lagrangian transforms as

\[
\mathcal{L} \to \mathcal{L} - e\overline{\Psi} \gamma^\mu (\partial_\mu \lambda(x)) \Psi'
\]

we see that we need to cancel the transformation of \( A_\mu \) with a transformation of \( \Psi \to \Psi' \). Let us try a local generalization of the global \( U(1) \) transformation

\[
\Psi \to e^{-ie\lambda(x)} \Psi
\]

The covariant derivative transforms as

\[
D_\mu \Psi = [\partial_\mu + ieA_\mu] \Psi \to [\partial_\mu + ieA'_\mu] \Psi' = [\partial_\mu + ie(A_\mu + \partial_\mu \lambda(x))] \Psi' = e^{-ie\lambda(x)} [\partial_\mu - ie\partial_\mu \lambda(x) + ie(A_\mu + \partial_\mu \lambda(x))] \Psi = e^{-ie\lambda(x)} D_\mu \Psi
\]

which shows that the covariant derivative has the convenient property that it only picks up a factor \( \exp(-ie\lambda(x)) \), with the derivative canceling the transformation of the gauge field. The whole Lagrangian Eq. (8.26) is then gauge invariant, since

\[
\overline{\Psi} \to \overline{\Psi} e^{+ie\lambda(x)}
\]

Let us summarize now the gauge symmetry of QED

\[
A_\mu(x) \to A_\mu(x) + \partial_\mu \lambda(x), \quad \Psi \to e^{-ie\lambda(x)} \Psi \tag{8.25}
\]

which leaves the Lagrangian and in particular \( i \overline{\Psi} \slashed{D} \Psi \) invariant.

### 8.5.2 Couple photons to charged scalars

We will write \( \varphi \) for the charged scalar field and not \( \psi \) as previously in order not to confuse it with the spinors.

How would we couple the charged scalar to a gauge boson? The linear coupling to the \( U(1) \) current would not work\(^6\) but we can use the covariant derivative \( D_\mu \)

\[
D_\mu = \partial_\mu + ieA_\mu
\]

and replace the \( \partial_\mu \) everywhere. This gives

\[
\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + (D_\mu \varphi)^* D^\mu \varphi - m^2 \varphi^* \varphi \tag{8.26}
\]

The ”covariant” is with respect to it’s behaviour under gauge transformations as we will see below.

\( \text{Try it!} \)
a simultaneous gauge transformation

\[ A_\mu(x) \rightarrow A_\mu(x) + \partial_\mu \lambda(x), \quad \varphi(x) \rightarrow e^{-ie\lambda(x)} \varphi(x) \quad (8.27) \]

This trick works for any theory. Replacing the derivatives by covariant derivatives in suitable representations works in almost any theory. This procedure is called minimal coupling.

### 8.6 The electric charge

Back to QED and the gauge-spinor coupling: the factor \( e \) has the interpretation of the coupling strength between the charged field and the photon. We know from electro-magnetism that \( j_0 \) is the charge density, as can be seen from the equations of motion, \( \partial_\mu F^{\mu\nu} = j^\nu \) or XXX. The total charge \( Q \) is then given by

\[ Q = e \int d^3x \bar{\Psi} \gamma^0 \Psi \]

Inserting the Heisenberg expansion of the free spinor, we get a similar expression as for the charged scalar field Eq. (5.60)

\[ Q = e \int \frac{d^3q}{(2\pi)^3} \sum_{r=1}^2 (b_q^\dagger b_q - c_q^\dagger c_q) \]

We find again that this is the number of particles minus the number of antiparticles. The value of the coupling constant in QED is usually written in terms of the fine-structure constant

\[ \alpha = \frac{e^2}{4\pi\hbar c} \approx \frac{1}{137} \quad (8.28) \]

which in natural units means for \( e = \sqrt{4\pi\alpha} \approx 0.3 \).

### 8.7 Discrete symmetries of QED

In addition to the Lorentz transformations which are continuously connected to the \( \mathbb{1} \), there are two other space-time operations which can be symmetries of the Lagrangian: parity \( \mathcal{P} \) and time-reversal \( \mathcal{T} \). Additionally, there can be a particle-antiparticle symmetry called charge conjugation \( \mathcal{C} \). We have already encountered parity as

\[ \mathcal{P} : (t, x) \rightarrow (t, -x) \]

which reverses handedness. Time reversal acts as

\[ \mathcal{T} : (t, x) \rightarrow (-t, x) \]

both leave the Minkowski interval \( x^2 = t^2 - x^2 \) invariant. Formally, we can say that the full Lorentz group breaks into four disconnected subsets. We call the continuous Lorentz transformation the proper, orthochronous Lorentz group \( \mathcal{L}_+^\uparrow \)

\[ \mathcal{L}_+^\uparrow \leftrightarrow \mathcal{P} \mathcal{L}_-^\uparrow \]

\[ \mathcal{L}_+^\downarrow \leftrightarrow \mathcal{T} \mathcal{L}_-^\downarrow \]

There is a subtlety here, in that these couplings are scale dependent

\[ \alpha = \alpha(\mu) \]

Once you include loop corrections, the value of the fine-structure constant grows logarithmically as the energy scale is increased. The value here is associated with the energy scale of the electron mass, so \( \alpha(\mu = m_e) \approx 1/137 \), however e.g. at the mass of the \( Z \) boson

\[ \alpha(\mu = M_Z) \approx 1/129 \]

The scale dependence is encoded in the renormalization group equation which you will discuss in the QFT lecture

\[ \frac{d\alpha(\mu)}{d\ln \mu} \equiv \beta(\epsilon) = \frac{\epsilon(\mu)^3}{12\pi^2} + \ldots \]

where just show the leading order dependence. The running of the coupling is a loop effect as you see from its \( \epsilon(\mu)^3 \) dependence.
Although a relativistic field theory must be invariant under $\mathcal{L}_+$, it is not necessarily invariant under $\mathcal{C}$, $\mathcal{P}$, or $\mathcal{T}$. The gravitational, electromagnetic, and the strong force respect all three, but weak interactions violate $\mathcal{C}$ and $\mathcal{P}$ separately, and certain suppressed processes violate $\mathcal{CP}$ and $\mathcal{T}$ violation. The combination $\mathcal{CPT}$ is observed in all processes so far in nature. There is also the famous CPT theorem which states that any Lorentz invariant local quantum field theory with a Hermitian Hamiltonian must have $\mathcal{CPT}$ symmetry. It is very difficult to write a consistent quantum field theory which would lead to observable $\mathcal{CPT}$ violation.

8.7.1 Parity

We have already discussed parity extensively. Let us just add the result for the gauge field and any four-vector

$$\mathcal{P} : A^\mu(x) = (A^0, A) \rightarrow (A^0, -A)$$

the Dirac spinor in the Weyl basis exchanges $\psi_L$ and $\psi_R$, whose representation is

$$\mathcal{P} : \Psi(x, t) \rightarrow \gamma^0 \Psi(-x, t)$$

8.7.2 Charge conjugation

We now know what the meaning of charge is in an interaction and hence we can talk about charge conjugation. We will try to flip the charge $e$ in

$$i\gamma^\mu(\partial_\mu + ieA_\mu)\Psi = 0$$

Start with the complex conjugation of the equation of motion

$$-i\gamma^{\mu*}(\partial_\mu - ieA_\mu)\Psi = 0$$

If we complex conjugate the Clifford algebra $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}1$, we see that $(-\gamma^{\mu*})$ also satisfies it. Therefore $(-\gamma^{\mu*})$ must be the gamma matrices expressed in a different basis, which means there exists a matrix $C\gamma^0$ such that

$$-\gamma^{\mu*} = (C\gamma^0)^{-1}\gamma^\mu C\gamma^0$$  \hfill (8.29)

We plug this into the conjugate equation of motion to find

$$i\gamma^\mu(\partial_\mu - ieA_\mu)\Psi_C = 0$$

with

$$\Psi_C \equiv C\gamma^0\Psi^*$$

Including $\gamma^0$ in the definition is convention.
Let us find the specific form of the charge conjugation matrix $C$. We can write Eq. (8.29) with $(\gamma^0)^2 = 1$ and so $(\gamma^0)^{-1} = \gamma^0$ as

$$C\gamma^0\gamma^\mu\gamma^0 C^{-1} = -\gamma^\mu$$

We know that the hermitian conjugate is $(\gamma^\mu)^\dagger = \gamma^0\gamma^\mu\gamma^0$ and taking the complex conjugate gives $(\gamma^\mu)^T = \gamma^0(\gamma^\mu)^*\gamma^0$ if $\gamma^0$ is real. Therefore we have

$$(\gamma^\mu)^T = -C^{-1}\gamma^\mu C$$

which explains why $C$ is defined with a $\gamma^0$ attached. In both bases (Weyl and Dirac from the Ex-sheet), $\gamma^2$ is the only imaginary matrix, so $\gamma^{\mu*} = \gamma^\mu$ for $\mu \neq 2$ and $\gamma^{2*} = -\gamma^2$. Then the defining equation Eq. (8.29) just says,

$$[C\gamma^0, \gamma^\mu] = 0, \quad \text{for } \mu \neq 2,$$

$$[C\gamma^0, \gamma^2] = 0$$

which means

$$C = -i\gamma^2\gamma^0$$

up to an arbitrary phase. So we find the simple relation for the charge conjugate spinor where we multiply by $(-i)$ for convenience

$$\Psi_C = -i\gamma^2\Psi^*$$

(8.30)

What does it mean in terms of the Weyl spinors? We find

$$\Psi_C = -i\gamma^2\left(\frac{\psi_L^2}{\psi_R^2}\right) = \left(\begin{array}{c} -i\sigma^2\psi_R^* \\ i\sigma^2\psi_L^* \end{array}\right)$$

This is interesting! So the charge conjugate of a left-handed field is right-handed and vice versa. In terms of the Lorentz representations, we have found an important relation, if

$$\psi_L : \left(\begin{array}{c} \frac{1}{2} \\ 0 \end{array}\right), \quad \text{then } i\sigma^2\psi_L^* : \left(\begin{array}{c} 0 \\ \frac{1}{2} \end{array}\right)$$

and vice versa. We can explicitly check that $\Psi_C$ transforms as a spinor with the known Lorentz transformation, Eq. (7.36), $\Psi \rightarrow e^{-\frac{i}{2}\omega_{\mu\nu}\sigma^{\mu\nu}}\Psi$ we have for the complex conjugate

$$\Psi \rightarrow e^{\frac{i}{2}\omega_{\mu\nu}(\sigma^{\mu\nu})^*}\Psi^*$$

hence

$$\Psi_C \rightarrow -i\gamma^2 e^{\frac{i}{2}\omega_{\mu\nu}(\sigma^{\mu\nu})^*}\Psi^*$$

$$= e^{-\frac{i}{2}\omega_{\mu\nu}(\sigma^{\mu\nu})}\Psi_C$$

where we have used the fact, that $[\gamma^2, \gamma^\mu\gamma^\nu] = 0$ for $\mu, \nu \neq 2$ and $[\gamma^2, \gamma^\nu\gamma^\nu] = 0$ for $\nu \neq 2$. This shows that $\Psi_C$ transforms like a regular Dirac spinor and confirms our statement about the representations of $\psi_L$ and $\psi_R$ being truly exchanging roles.

Recall, that

$$\gamma^\text{Weyl} = \left(\begin{array}{cc} 0 & \sigma^2 \\ -\sigma^2 & 0 \end{array}\right)$$

with

$$\sigma^2 = \left(\begin{array}{cc} 0 & -i \\ i & 0 \end{array}\right)$$

We choose $-i$ for convenience.

Experimentally, we know that the neutrino is left-handed. Thus, we are now able to predict that the anti-neutrino is right-handed!

This is due to the fact the representations of $SU(2)$ are pseudo-real. A representation and its complex conjugate are related to each other by a simple transformation $S$, e.g. for the complex doublet $2$

$$\Psi^* = S^{-1}2S$$

in this case with $\varepsilon_{ij} = i\sigma^3_{ij}$ it is just

$$2^*_i = \varepsilon_{ij}2_j$$

Since for $\mu, \nu \neq 2$

$$(\sigma_{\mu\nu})^* = \left(\frac{1}{2}[\gamma_\mu, \gamma_\nu]\right)^*$$

$$= -\frac{1}{2}[\gamma_\mu, \gamma_\nu]$$

$$= -\frac{1}{2}[\gamma_\mu, \gamma_\nu]$$

$$= -\sigma_{\mu\nu}$$

since these $\gamma$ matrices are real (only $\gamma^3$ is purely imaginary). If one of the indices $\mu = 2$ or $\nu = 2$, then e.g.

$$(\sigma_{2\nu})^* = -\frac{i}{2}(-\gamma_2, (\pm)\gamma_\nu)$$
8.7.3 Majorana neutrino

Because $\Psi_C$ transforms as a spinor, Majorana understood that Lorentz invariance no only allows the Dirac equation $i\partial\Psi = m\Psi$ but also the Majorana equation

$$i\partial\Psi = m\Psi_C$$

(8.31)

This type of mass term is called a Majorana mass. We can obtain it from the Lagrangian

$$\mathcal{L} = \overline{\Psi} i\partial\Psi - \frac{1}{2} m \left( \Psi^T C \Psi + \Psi C \Psi^T \right)$$

Since $\Psi$ and $\Psi_C$ carry opposite charge, we can use the Majorana equation, unlike the Dirac equation, only for electrically neutral fields. You can also see this from the Lagrangian: the mass term breaks the $U(1)$ symmetry $\Psi \to e^{i\alpha} \Psi$. We can therefore also not couple it to gauge fields. How does it look in components? The mass contains terms like

$$i(\sigma_2)_{ij} = \varepsilon_{ij}$$

we have $C_{\alpha\beta} = -C_{\beta\alpha}$ this seems to indicate that the term has to vanish if the spinor fields $\Psi_\alpha$ commute. In your QFT lecture, you will learn that spinors $\Psi$ have to be treated as anticommuting Grassmannian numbers,

$$\Psi_1 \Psi_2 = -\Psi_2 \Psi_1$$

which we have already (secretly) realised when defining time-ordered or normal ordered products of Fermion fields.

In terms of the Weyl two-spinors $\Psi = (\psi_L, \psi_R)$ the mass term looks like

$$\frac{1}{2} m \Psi^T C \Psi + \text{h.c.} = -\frac{1}{2} m (\psi_L^T i\sigma_2 \psi_L - m \psi_R^T i\sigma_2 \psi_R + \text{h.c.})$$

We see that we could have been more economical and started with just one Weyl fermion, $\psi_L$ since the $\psi_L$ and $\psi_R$ fields are now not mixing through the mass term. A minimal real Majorana field would be

$$\mathcal{L} = i\partial_\mu \gamma^\mu \psi_L + m (\psi_L^T i\sigma_2 \psi_L + \text{h.c.})$$

(8.32)

which brings us back to the initial mystery of how to give a single Weyl fermion a Lorentz-invariant mass term. All that we needed to save the naive guess in Eq. (7.7), was an anti-symmetric matrix.

$$i\sigma_\mu \partial_\mu \psi_L - m \psi_L^\dagger i\sigma_2 = 0$$

As we know, they are irreducible representations of the Lorentz group and will be separately Lorentz covariant.
What happens to spinor bilinears under charge conjugation? You can easily check that

$$\mathcal{C} : \bar{C} \gamma^\mu \Psi C \rightarrow -\bar{C} \gamma^\mu \Psi$$

This means that the current $j^\mu$ of QED changes sign.

### 8.7.4 Furry’s theorem

We can start with the photon one-point function. We note that the external photon must be attached to a QED vertex. Neglecting the external propagator, the amplitude is hence

$$\langle ...j'(x) = -ie \int d^4x e^{-iq\cdot x} \langle \Omega | T j_\mu(x) | \Omega \rangle$$

where $j^\mu = \bar{C} \gamma^\mu \Psi$ is the electro-magnetic current operator. The vacuum expectation value of $j^\mu$ has to vanish because of Lorentz-invariance: otherwise $j^\mu$ would be a preferred direction in Minkowski space.

The photon one-point function has to vanish also for other reasons, namely charge conjugation invariance. Since $\mathcal{C}$ is symmetry of QED, we know that the vacuum $|0\rangle$ and the vacuum of the interacting theory

$$C|\Omega\rangle = |\Omega\rangle$$

are $\mathcal{C}$ invariant. A diagram with $n$ external photons (and no external fermions) is proportional to

$$\langle \Omega | T j_{\mu_1}(x_1) \cdots j_{\mu_n}(x_n) | \Omega \rangle$$

If $n$ is odd then

$$\langle \Omega | T j_{\mu_1}(x_1) \cdots j_{\mu_n}(x_n) | \Omega \rangle = \langle \Omega | T C j_{\mu_1}(x_1) C C \cdots C j_{\mu_n}(x_n) C C | \Omega \rangle$$

$$= \langle \Omega | T (Cj_{\mu_1}(x_1)C)(C \cdots C)(Cj_{\mu_n}(x_n)C) | \Omega \rangle$$

$$= (-1)^n \langle \Omega | T j_{\mu_1}(x_1) \cdots j_{\mu_n}(x_n) | \Omega \rangle$$

This means that the a pure photon $n$-point function vanishes, for $n$ odd. This is known as Furry’s theorem.

### 8.8 Time reversal

#### 8.8.1 Quantum mechanics

We will now complete our discussion of the discrete symmetries of QED and discuss time reversal invariance. Time reversal symmetry is a bit subtle as you might already know from your quantum mechanics course. Let us investigate this first using the Schrödinger equation of QM

$$\frac{\partial}{\partial t} \psi(t) = H \psi(t)$$

We are suppressing the trivial $x$ dependence in this section.
with for concreteness think $H = -\frac{1}{2m} \nabla^2 - V(x)$. We now consider
the time reversal transformation

$$t \rightarrow t' = -t$$

Our goal is now to find $\psi'(t')$ which satisfies

$$i \frac{\partial}{\partial t'} \psi(t') = H \psi(t')$$

We write for the transformed field

$$\psi'(t') = T \psi(t)$$

where $T$ is an operator which we will determine in the following. Let us plug this into Eq. (8.35), we obtain

$$i \frac{\partial}{\partial(-t)} T \psi(t) = HT \psi(t)$$

$$T^{-1} i \frac{\partial}{\partial(-t)} T \psi(t) = T^{-1} HT \psi(t)$$

Since $H$ is time-independent by assumption, the time-inversion operator has no effect and

$$T^{-1} H = HT^{-1},$$

which results in

$$T^{-1}(-i)T \frac{\partial}{\partial t} \psi(t) = H \psi(t)$$

Comparing to Eq. (8.34), we are not forced to conclude

$$T^{-1}(-i)T = i$$

We can define

$$T \equiv UK$$

where $K$ complex conjugates everything to the right. Applied to Eq. (8.36) we get with $T^{-1} = KU^{-1}$

$$KU^{-1}(-i)U K = U^{-1}(-i)U^* K^2 = U^{-1}(-i)U^*$$

We see that it does the job, if $U^{-1}iU = i$, that is, if $U^{-1}$ is just an ordinary (unitary) operator that leaves $i$ be.\(^{10}\) The contribution of $K$ to $T$ makes it antiunitary.

### 8.8.2 Scalar Field

Let us check how $T$ acts on a scalar state solving the free Klein-Gordon equation $\phi(t) = e^{-iE x} = e^{i(k \cdot x - Et)}$. We have

$$\phi'(t') = T \phi(t) = UK \phi(t) = U \phi^*(t) = U e^{-i(k \cdot x - Et)}$$

Since $\phi$ has just one component, $U$ is a trivial phase-factor.\(^{11}\) We can rewrite this as

$$\phi'(t) = e^{-i(k \cdot x + Et)} = e^{i(-k \cdot x - Et)}$$

from which we see that $\phi'$ describes a wave moving in the opposite direction. The energy is still positive since $\psi \propto e^{-iEt}$. We see that two time-reversals are equal to the identity

$$T^2 = UKUK = UU^* K^2 = +1$$
8.8.3 Spinor Field

Now we need to consider how $T$ acts on spinors. The Dirac equation is after multiplying by $\gamma^0$ from the left

$$i \frac{\partial}{\partial t} \Psi(t) = H \Psi(t)$$

with $H = -i\gamma^0\gamma^i \partial_i + \gamma^0 m$. As above, we want

$$i \frac{\partial}{\partial t'} \Psi'(t') = H \Psi(t')$$

with

$$\Psi'(t') = T\Psi(t)$$

Following our discussion above, we can make similar arguments if $T^{-1}HT = H$ or

$$KU^{-1}HU = H$$

which means we require

$$KU^{-1}\gamma^0UK = \gamma^0$$

and

$$KU^{-1}(i\gamma^0\gamma^i)UK = i\gamma^0\gamma^i$$

If we move $K$ to the RHS, we see that we can determine $U$ with

$$U^{-1}\gamma^0U = \gamma^0, \quad \text{and} \quad U^{-1}\gamma^iU = -\gamma^i$$

Restricting ourselves to the Weyl (and Dirac) basis, we know that $\gamma^2$ is the only imaginary matrix. So we need to find a $U$ which flips $\gamma^1$ and $\gamma^3$ but not $\gamma^0$ and $\gamma^2$? This choice works\(^{12}\)

$$U = \gamma^1\gamma^3$$

and so

$$\Psi(t, x) \rightarrow \gamma^1\gamma^3\Psi(-t, x) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \Psi(-t, x)$$

Thus $T$ flips the spin of particles.

8.9 Electron scattering and the Coulomb potential

We describe electron scattering $e^- e^- \rightarrow e^- e^-$ with the two tree-level Feynman diagrams as

\(^{12}\) Again, with an arbitrary phase which we set to 1.
Summary of $C$, $P$, and $T$

The transformation properties of the various fermion bilinears under $C$, $P$, and $T$ are summarized in the table below. Here we use the shorthand $(-1)^{\mu} \equiv 1$ for $\mu = 0$ and $(-1)^{\mu} \equiv -1$ for $\mu = 1, 2, 3$.

<table>
<thead>
<tr>
<th></th>
<th>$\bar{\psi}\psi$</th>
<th>$i\bar{\psi}\gamma^5\psi$</th>
<th>$\bar{\psi}\gamma^\mu\psi$</th>
<th>$\bar{\psi}\gamma^\mu\gamma^5\psi$</th>
<th>$\bar{\psi}\sigma^{\mu\nu}\psi$</th>
<th>$\partial_\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P$</td>
<td>+1</td>
<td>-1</td>
<td>$(-1)^\mu$</td>
<td>$(-1)^\mu(-1)^\nu$</td>
<td>$(-1)^\mu$</td>
<td>$\partial_\mu$</td>
</tr>
<tr>
<td>$T$</td>
<td>+1</td>
<td>-1</td>
<td>$(-1)^\mu$</td>
<td>$(-1)^\mu(-1)^\nu$</td>
<td>$(-1)^\mu$</td>
<td>$\partial_\mu$</td>
</tr>
<tr>
<td>$C$</td>
<td>+1</td>
<td>+1</td>
<td>-1</td>
<td>+1</td>
<td>-1</td>
<td>+1</td>
</tr>
<tr>
<td>$CPT$</td>
<td>+1</td>
<td>+1</td>
<td>-1</td>
<td>-1</td>
<td>+1</td>
<td>-1</td>
</tr>
</tbody>
</table>

\[
\begin{align*}
    iA &= (-ig)^2 \left[ \bar{u}(q')\gamma^\mu u(q) \frac{(-i\eta_{\mu\nu})}{(q-q')^2 - \mu^2} \bar{u}(p')\gamma^\nu u(p) \\
    &\quad - \bar{u}(p')\gamma^\mu u(q) \frac{-i\eta_{\mu\nu}}{(p'-q')^2 - \mu^2} \bar{u}(q')\gamma^\nu u(p) \right] \quad (8.38)
\end{align*}
\]

Compared to the scalar mediator we seem to find an overall minus sign, but remember that the expression is really positive for $\mu, \nu = 1, 2, 3$.

8.9.1 Coulomb potential

We take the first diagram of $e^- e^- \rightarrow e^- e^-$ and calculate the non-relativistic limit. The steps proceed as in Sec. 6.7.1 and Sec. 7.8.4 for the Yukawa potential. The only difference is that we now need to evaluate $\bar{u}\gamma^\mu u$ in Eq. (8.38) using the non-relativistic spinor

\[
u(p) \rightarrow \sqrt{m} \begin{pmatrix} \xi \\ \xi \end{pmatrix}
\]

We see that

\[
\bar{u}^r(p)\gamma^0 u^s(q) \rightarrow 2m\delta^{rs}
\]

\[
\bar{u}^r(p)\gamma^i u^s(q) \rightarrow 0
\]

Comparing the scattering amplitude in the non-relativistic limit to the quantum mechanical expression again, we find an effective potential between two electrons

\[
U(r) = +e^2 \int \frac{d^3p}{(2\pi)^3} \frac{e^{ip\cdot r}}{|r|^2} = \frac{e^2}{4\pi r}
\]

Which is the familiar repulsive Coulomb potential between two electrons. What about $e^+ e^- \rightarrow e^- e^+$ scattering? The relevant diagram in the non-relativistic limit is
which gives

\[ iA = -(-ig)^2 \left[ \bar{\psi}(q')\gamma^\mu v(q) \frac{(-i\eta_{\mu\nu})}{(q - q')^2} \bar{u}(p')\gamma^\nu u(p) \right] \]  

(8.39)

The overall + sign comes from treating the fermions correctly: we saw the same minus sign when studying scattering in Yukawa theory. The non-relativistic limit is now different however, since \( \bar{\psi}\gamma^0 v \rightarrow 2m \) which gives us

\[ U(r) = -e^2 \int \frac{d^3p}{(2\pi)^3} \frac{e^{ip\cdot r}}{|p|^2} = -\frac{e^2}{4\pi r} \]

As expected, we find an attractive force between an electron and a positron. We can see the same effect without all the spinor-ology in scalar QED. The interaction is \( (D_\mu \phi)^* D^\mu \phi \) and so the relevant term for the scalar-photon interaction is

\[ -ieA_\mu(\phi^* \partial^\mu \phi - \phi \partial_\mu \phi^*) + e^2 A_\mu A^\mu \]

the two vertices are

\[ = -ie(p + q)_\mu, \quad = +2ie^2 \eta_{\mu\nu} \]

(8.40)

The momentum dependence comes from the \( \partial^\mu \) term and the factor of 2 is because of the identical particles \( A_\mu A^\mu \) appearing in the Lagrangian.

We can see that the difference in sign comes from the \( A_0 \) piece of the propagator \( -i\eta_{\mu\nu}/p^2 \), we have in the non-relativistic limit of scalar \( e^- e^- \rightarrow e^- e^- \) scattering

\[ = -i\eta_{\mu\nu}(-ie)^2 \frac{(p + p')^\mu(q + q')_\nu}{(p' - p)^2} \rightarrow -i(-ie)^2 \frac{(2m)^2}{-(p - p')^2} \]

where the numerator selects the \( A^0 \) piece since \( (p + p')^\mu(q + q')_\mu \approx (p + p')^0(q + q')^0 \approx (2m)^2 \). The Coulomb potential for spin 0 particles is again repulsive. If switch to scalar \( e^- e^- \rightarrow e^- e^- \) scattering, one of the arrows on the legs changes direction and the amplitude picks up an additional minus sign because since the charge arrows are correlated with momentum arrows. This leads as expected to an attractive potential

Compare to the Yukawa potential

where the limit is \( \bar{\psi}v \rightarrow -2m \).
9

Outlook

If you found this interesting you should continue and attend the QFT course in the next semester. You have yet to encounter many of the truly awesome and deep insights that quantum field theory contains. It is the language in which the laws of Nature are written. As Sidney Coleman said: "Not only God knows, I know, and by the end of the semester, you will know."