

Quantum Field Theory I

Arthur Hebecker

Institute for Theoretical Physics, Heidelberg University

Originally type-set mostly from my hand-written notes in the Winter Term 2015/16 by **Robert Ott, Martin Pauly, Fabio Schlindwein, and Jan Schneider**. To be further edited as the 2021/22 lecture course proceeds....

Contents

1	Introduction	3
1.1	Motivation	3
1.2	Symmetries (Poincaré transformations)	4
1.3	Symmetries acting on fields	5
2	Free Scalar Field	8
2.1	Classical theory - Lagrangian formulation	8
2.2	Classical theory - Hamiltonian formulation	10
2.3	Quantisation: real scalar	12
2.4	Quantisation: complex scalar	16
3	Noether's theorem	17
3.1	Formulation and derivation in field theory	17
3.2	Energy-Momentum-Conservation	19
3.3	U(1)-symmetry and charge of a complex scalar	21
4	Heisenberg picture, causality, covariance	22
4.1	Heisenberg picture	22
4.2	Causality	23
4.3	Covariance	23
5	Perturbation theory: first steps in a naive, leading-order approach	26
5.1	S-Matrix	26
5.2	Scattering cross section	31
5.3	2-particle phase-space and a simple example	36
6	LSZ-Formalism	37
6.1	Spectral density and Z-factors	37
6.2	LSZ reduction formula	40
6.3	Calculating time-ordered correlation functions	43

7	Wick-Theorem and Feynman Rules	45
7.1	Time ordering and normal ordering	45
7.2	Wick theorem	47
7.3	The Feynman propagator	48
7.4	Feynman rules	48
7.5	Feynman Rules in Momentum Space	52
7.6	Calculating the Z-Factor and the Physical Mass	56
7.7	Feynman rules for scattering amplitudes	58
8	The Electromagnetic Field	60
8.1	Gauge invariance	60
8.2	Gupta-Bleuler Quantization	64
8.3	Photon Propagator	69
8.4	Feynman rules for scalar QED	70
9	Spinors	73
9.1	Fields and representations	73
9.2	Remarks on Lie Groups & Lie Algebras	74
9.3	The spinor representation of $SO(1,3)$	77
9.4	Invariants involving spinors and lagrangian equations of motion	82
9.5	Solutions of the Dirac equation	84
10	Quantization of Spinors	86
10.1	Hamiltonian	86
10.2	Quantization attempts with commutators	87
10.3	Quantization with anticommutators	88
10.4	Time ordering, Green's functions, Dirac propagators	91
10.5	$U(1)$ -Symmetry of the Dirac lagrangian	94
10.6	Yukawa theory	94
11	Quantum Electrodynamics	95
11.1	Lagrangian	95
11.2	Deriving the Feynman rules	96
11.3	Summary of QED Feynman rules and simple examples	100
11.4	Elementary processes	101
12	Renormalization	108
12.1	Concept	108
12.2	Renormalization conditions	109
12.2.1	Mass and field (re)normalisation of a scalar	109
12.2.2	Mass and field normalization of the electron	111
12.2.3	Field normalization of the photon	112
12.2.4	Vertex normalization	114
12.3	Vacuum polarisation in dimensional regularisation	116
12.4	QED Beta-function	119
12.5	The Ward-Takahashi identity	122
12.6	Sketch of an operator derivation of the Ward Takahashi identity	125

13 Non-abelian Gauge Theory and Standard Model	125
13.1 Non-abelian gauge theory	125
13.2 Standard Model	127

1 Introduction

Our main reference is the textbook by Peskin and Schroeder [1]. Further literature will be given as we go along. See also the Web page of the course.

1.1 Motivation

Our world is fundamentally quantum mechanical. Classicality arises only approximately, in a particular limit. Moreover, fields are part of our reality – e.g. the electromagnetic field strength and potential $\vec{E}(\vec{x}, t)$, $\vec{B}(\vec{x}, t)$, $A_\mu(\vec{x}, t)$. Thus, the fundamental description of our world needs a quantum theory of fields.

Clearly, the photon and its interactions will be part of this theory. It is less obvious but nevertheless true that also the electron will also emerge as a quantum of an appropriate field. In fact, all particles of the “Standard Model” are quanta of fields as above. At energies below the Planck scale $M_P = 1/\sqrt{G_N}$, even the graviton falls into that scheme.¹

Thus, quantum field theory or QFT is **the fundamental theory of this world** (leaving aside conjectural theories or quantum gravity models, like for example string theory). Quantum mechanics is its non-relativistic limit. QFT is also the most precisely tested theory we have.

In addition, so-called **effective fields** are central in Condensed Matter Theory (CMT), see e.g. Fig. 1.

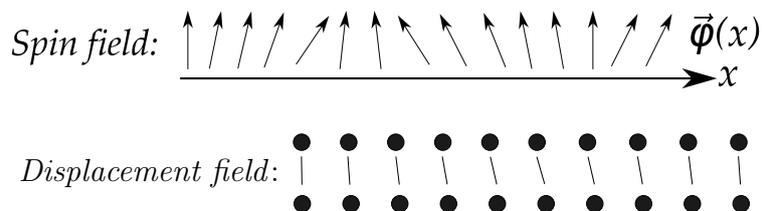


Figure 1: Examples of fields relevant in condensed matter.

Most dynamics of these fields is in the quantum regime, thus we once again need a QFT description. As a result, QFT is *the* modern language not only of particle physics but also of CMT, including related popular research areas like ultracold atoms. The role of QFT is not diminished by the fact that, in these contexts, it represents only an effective description of what is fundamentally many-body quantum mechanics.

¹ We will try to always work in units where $\hbar = c = 1$. Moreover, in the modern research literature one mostly uses the so-called reduced Planck mass $\overline{M}_P = M_P/\sqrt{8\pi} \simeq 2.4 \times 10^{18}$ GeV. The ‘bar’ over M_P is often dropped, so one always needs to be careful which one is meant.

1.2 Symmetries (Poincaré transformations)

As in all of physics, the symmetries of the system we study will be essential. Our stage is space-time: $\mathbb{R}^4 \ni (t, \vec{x}) = \{x^\mu\}$. To describe its symmetries, let us first recall the familiar symmetries of space by itself: $\mathbb{R}^3 \ni \vec{x} = \{x^i\}$; $i = 1, 2, 3$. They are given by the group of translations and rotations:²

$$\vec{x} \longrightarrow \vec{x}' \quad \text{with} \quad x'^i = R^i_j x^j + d^i. \quad (1.1)$$

Let us recall how one determines which matrices R are allowed: The key condition is that the length of vectors,

$$|\vec{x}|^2 = \sum_{i=1}^3 (x^i)^2 = x^i x^j \delta_{ij}, \quad (1.2)$$

should not change. Here δ_{ij} is the euclidean metric on \mathbb{R}^3 . In other words, we demand

$$x'^i x'^j \delta_{ij} = x^i x^j \delta_{ij} \quad (1.3)$$

or

$$\delta_{ij} R^i_k x^k R^j_l x^l = \delta_{ij} x^i x^j \quad \forall x \in \mathbb{R}^3 \quad (1.4)$$

or

$$\delta_{ij} R^i_k R^j_l = \delta_{kl} \quad \Leftrightarrow \quad R R^\top = \mathbb{1} \quad \Leftrightarrow \quad R \in O(3). \quad (1.5)$$

In summary, physical space is \mathbb{R}^3 with euclidean metric. Its symmetries are translations and rotations. The latter are defined as linear transformations leaving the metric invariant.

The generalization to space-time and Poincaré invariance is straightforward:

$$\mathbb{R}^3 \ni \{x^i\} \quad \longrightarrow \quad \mathbb{R}^4 \ni \{x^\mu\} = (t, \vec{x}) \quad (1.6)$$

$$\delta_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}_{ij} \quad \longrightarrow \quad \eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}_{\mu\nu}. \quad (1.7)$$

Here overall sign of $\eta_{\mu\nu}$ is pure convention. The relative sign between time part and space part of η is instead deep physical reality. If you wish it is observational data summarized in the fact that special relativity governs our world (as long as general-relativistic effects are negligible).

The length-squared of a vector generalizes as

$$|\vec{x}|^2 = x^i x^j \delta_{ij} \quad \longrightarrow \quad x^2 = x^\mu x^\nu \eta_{\mu\nu} = t^2 - \vec{x}^2, \quad (1.8)$$

where we recall that $c = 1$ throughout this course. If x characterizes the separation of two points, then one refers to $x^2 > 0$ and $x^2 < 0$ as time-like and space-like separation respectively. The symmetry group generalizes according to

$$R \in O(3) \text{ if } \delta_{ij} R^i_k R^j_l = \delta_{kl} \quad \longrightarrow \quad \Lambda \in O(1,3) \text{ if } \eta_{\mu\nu} \Lambda^\mu_\rho \Lambda^\nu_\sigma = \eta_{\rho\sigma}. \quad (1.9)$$

² Actually, rotations and reflections, but it will be too painful to always mention the latter explicitly.

With this, **Poincaré transformations** are defined by

$$x^\mu \longrightarrow x'^\mu = \Lambda^\mu_\nu x^\nu + d^\mu . \quad (1.10)$$

In other words, the Poincare group is the group of pairs (Λ, d) with $\Lambda \in O(1, 3)$, $d \in \mathbb{R}^4$ and the composition law

$$(\Lambda_1, d_1) \cdot (\Lambda_2, d_2) = (\Lambda_1 \cdot \Lambda_2, \Lambda_1 \cdot d_2 + d_1) . \quad (1.11)$$

This will be our most important symmetry. We will call \mathbb{R}^4 with this symmetry $\mathbb{R}^{1,3}$ or **Minkowski space**. The subgroup respecting the origin are known as the **Lorentz-transformations** Λ . Clearly, rotations may be viewed as are a subgroup of Lorentz transformations. In other words, we may call

$$\Lambda = \begin{pmatrix} 1 & \vec{0}^\top \\ \vec{0} & R \end{pmatrix} \quad \text{with} \quad R \in SO(3) \quad (1.12)$$

a rotation in $O(1, 3)$.

The subgroup of **special Lorentz-transformations** is defined by demanding

$$\det(\Lambda) = 1 \quad \text{and} \quad \Lambda^0_0 > 0 . \quad (1.13)$$

The also form the **identity component** $SO^+(1, 3) \subset O(1, 3)$, i.e. those elements which can be continuously connected to the identity.

For $\mathbb{R}^{1,1}$, Λ is obviously just a 2×2 Matrix. In this case it is easy to be very explicit and convince oneself that

$$\Lambda = \begin{pmatrix} \cosh(\alpha) & \sinh(\alpha) \\ \sinh(\alpha) & \cosh(\alpha) \end{pmatrix} \in SO(1, 1) \quad (1.14)$$

is the general group element. The corresponding transformation reads

$$\begin{pmatrix} t \\ x \end{pmatrix} \mapsto \Lambda \begin{pmatrix} t \\ x \end{pmatrix} = \begin{pmatrix} t \cdot \cosh(\alpha) + x \cdot \sinh(\alpha) \\ x \cdot \cosh(\alpha) + t \cdot \sinh(\alpha) \end{pmatrix} . \quad (1.15)$$

This is obviously a boost with $\beta = v/c = v$ and $\cosh \alpha = 1/\sqrt{1 - \beta^2}$ as well as $\sinh \alpha = \beta/\sqrt{1 - \beta^2}$.

1.3 Symmetries acting on fields

Consider some scalar field configuration, including its classical evolution in time. Mathematically, this is given by a function

$$\varphi : \mathbb{R}^4 \rightarrow \mathbb{R} \quad ; \quad x \mapsto \varphi(x) . \quad (1.16)$$

For simplicity, let us first replace \mathbb{R}^4 by \mathbb{R} (i.e. imagine a 1d world, without time). Now, consider a field configuration localized near zero and apply a translation by d , cf. Fig. 2. Note that we are taking the ‘active point of view’ concerning symmetries: We are not considering coordinate transformation but instead an actual change of the physical field configuration.

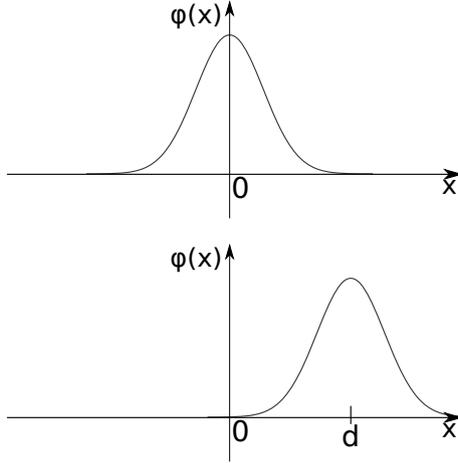


Figure 2: Translation applied to a classical field configuration.

Concretely, we want to focus on a translation

$$x \rightarrow x' = x + d \quad (1.17)$$

and determine the corresponding transformation of a generic field configuration,

$$\varphi \rightarrow \varphi'. \quad (1.18)$$

The defining property is

$$\varphi'(x') = \varphi(x), \quad (1.19)$$

which is the same as

$$\varphi'(x + d) = \varphi(x) \quad \text{or} \quad \varphi'(x) = \varphi(x - d). \quad (1.20)$$

This makes sense also intuitively: if φ had its maximum at $x = 0$, then φ' has its maximum at $x = d$. Thus, φ' is defined by applying the inverse transformation to the argument.

This story works completely analogously for the case we are really interested in: Poincaré transformations in $\mathbb{R}^{1,3}$. For

$$(\Lambda, d) : \varphi \mapsto \varphi' \quad (1.21)$$

we have:

$$\varphi'(x') = \varphi(x) \quad \text{or} \quad \varphi'(\Lambda x + d) = \varphi(x). \quad (1.22)$$

Now we define y by $x = \Lambda^{-1}y$ and find

$$\varphi'(y + d) = \varphi(\Lambda^{-1}y). \quad (1.23)$$

After a further change of variables, $y = x - d$ (this x is unrelated to the x used initially), we eventually find

$$\varphi'(x) = \varphi(\Lambda^{-1}(x - d)). \quad (1.24)$$

As before, the field transforms by applying the inverse transformation to the argument.

Next, let us look at the transformation of a vector built out of derivatives of fields:

$$\{\partial_\mu\varphi\} \equiv \left\{ \frac{\partial}{\partial x^\mu}\varphi \right\} \equiv \left\{ \frac{\partial}{\partial x^0}\varphi(x^0, \dots, x^3), \frac{\partial}{\partial x^1}\varphi(x^0, \dots, x^3), \dots, \dots \right\}. \quad (1.25)$$

We want to calculate

$$\partial_\mu\varphi' = \frac{\partial}{\partial x^\mu}\varphi(\Lambda^{-1}x), \quad (1.26)$$

where we have set $d = 0$ since it will not be essential for us. Introducing $y = \Lambda^{-1}x$ we find

$$\partial_\mu\varphi' = \frac{\partial}{\partial x^\mu}\varphi(y(x)) = \frac{\partial y^\nu}{\partial x^\mu} \frac{\partial}{\partial y^\nu}\varphi(y) = \frac{\partial}{\partial x^\mu} ((\Lambda^{-1})^\nu{}_\rho x^\rho) \partial_\nu(\varphi(y)) = (\Lambda^{-1})^\nu{}_\mu (\partial_\nu\varphi)(y). \quad (1.27)$$

We would like to say that our vector transforms, in addition to the familiar Λ^{-1} acting on the argument, with an extra Λ acting on the vector index. But the index contraction is not quite right.

To resolve this, let us make a small mathematical excursion into **dual vector spaces** and the **inverse metric**:

Let $x^\mu, \mu = 0, \dots, 3$ be an element of $\mathbb{R}^{1,3} = V$. Elements of V^* are denoted by $y_\mu, \mu = 0, \dots, 3$ such that $xy \equiv x^\mu y_\mu$. The metric η provides a natural map $V \rightarrow V^* : x^\mu \mapsto \eta_{\mu\nu}x^\nu$. This map and its inverse are often referred to as ‘lowering/raising indices’. For raising an index, the inverse metric $\eta^{\mu\nu} = \text{diag}(1, -1, -1, -1)$ is used: $x_\mu \mapsto \eta^{\mu\nu}x_\nu$. We have $\eta^{\mu\nu}\eta_{\nu\rho} = \delta^\mu{}_\rho \equiv \eta^\mu{}_\rho$.

Now, we already know that $\eta_{\mu\nu}\Lambda^\mu{}_\rho\Lambda^\nu{}_\sigma = \eta_{\rho\sigma}$. Using matrix notation, this can be rewritten as

$$\Lambda^T\eta\Lambda = \eta \quad \text{and hence} \quad \eta^{-1}\Lambda^T\eta = \Lambda^{-1}. \quad (1.28)$$

In index notation, this takes the form

$$(\Lambda^{-1})^\mu{}_\nu = \eta^{\mu\rho}(\Lambda^T)_\rho{}^\sigma\eta_{\sigma\nu} = \eta^{\mu\rho}\Lambda^\sigma{}_\rho\eta_{\sigma\nu} \equiv \Lambda_\nu{}^\mu. \quad (1.29)$$

Here, in the last expression, we introduced $\Lambda_\nu{}^\mu$ as a ‘ Λ -matrix with lowered first and raised second index.’ We have learned that this object is identical to the inverse Λ -matrix.

Equation (1.27) now takes the form

$$\partial_\mu\varphi'(x) = \Lambda_\mu{}^\nu\partial_\nu\varphi(\Lambda^{-1}x). \quad (1.30)$$

Thus, the derivative of a field transforms (in addition to the usual transformation through its argument) as an element of V^* , i.e. analogously to $x_\mu \rightarrow \Lambda_\mu{}^\nu x_\nu$. This is in contrast to how an element of V transforms: $x^\mu \rightarrow \Lambda^\mu{}_\nu x^\nu$.

2 Free Scalar Field

2.1 Classical theory - Lagrangian formulation

As should be known from the corresponding classical physics course, electrodynamics may be defined by the action

$$S = \frac{1}{4} \int d^4x F_{\mu\nu} F^{\mu\nu} \quad \text{with} \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \quad (2.1)$$

While the quantisation of this theory, which is of overwhelming practical importance, represents one of the main goals of our course, it is not a good starting point. The reason is that our task is complicated by A_μ having 4 components and by the gauge invariance under $A_\mu \rightarrow A_\mu + \partial_\mu \chi$.

Hence, we will first study a toy model which is obtained roughly by replacing the so-called **vector field** $A_\mu(x)$ with the **scalar field** $\varphi(x)$. This theory is also directly relevant to the real world as it describes (in slightly extended forms) the Higgs particle as well as, for example, pions.

We formulate the theory in analogy to mechanics. In **(1-dimensional) mechanics**, the key dynamical variable q is characterised by the function

$$q : t \mapsto q(t). \quad (2.2)$$

Its dynamics is governed by an action

$$S = S[q] = \int dt L(q, \dot{q}). \quad (2.3)$$

In **scalar field theory**, the analogous objects are the field

$$\varphi : x \mapsto \varphi(x) \quad , \quad x \in \mathbb{R}^4 \quad (2.4)$$

and the action

$$S = S[\varphi] = \int dt L[\varphi(t, \vec{x}), \dot{\varphi}(t, \vec{x})]. \quad (2.5)$$

We assume that L is **local** in \vec{x} , i.e.

$$L = \int d^3x \mathcal{L}(\varphi(x), \dot{\varphi}(x), \vec{\nabla}\varphi(x), \vec{\nabla}\dot{\varphi}(x), \dots), \quad (2.6)$$

with only **finitely many higher derivatives** appearing. (We recall that x is equivalent to (t, \vec{x}) .) An equivalent formulation is

$$S = \int d^4x \mathcal{L}(\varphi, \dot{\varphi}, \vec{\nabla}\varphi, \vec{\nabla}\dot{\varphi}, \dots) \quad (2.7)$$

where, again, the ellipsis stands for only finitely many higher spatial derivatives of φ or $\dot{\varphi}$. We call this ‘locality’ because a ‘non-local’ expression like $(\varphi(t, \vec{x}) - \varphi(t, \vec{x} + \vec{\delta}))$ would, if Taylor-expanded, involve infinitely many higher derivatives. The quantity \mathcal{L} is known as the **lagrangian density**, not to be confused with the

lagrangian L . However, since in field theory \mathcal{L} appears much more frequently than L , everybody calls \mathcal{L} simply the lagrangian, returning to the proper naming only when both quantities are used together.

Concerning derivatives w.r.t. the time variable t , we also follow the example of classical mechanics and demand that only **the first derivative** is present. One could roughly say that we generalise the locality of $S[q]$ of mechanics in t to locality of $S[\varphi]$ in $x \in \mathbb{R}^{1,3}$.

Next we define $V(\varphi) \equiv -\mathcal{L}(\varphi = \text{const.})$ and assume that V has a minimum at φ_0 . Without loss of generality, let $\varphi_0 = 0$ & $V(\varphi_0) = 0$ such that

$$V(\varphi) = \frac{1}{2}m^2\varphi^2 + \dots \quad (2.8)$$

Here we have disregarded, for the moment, higher terms in the Taylor series since our interest is in small fluctuations around the ground state $\phi = 0$ (or equivalently small excitations of the vacuum/ground state).

We now have

$$\mathcal{L} = T - V, \quad (2.9)$$

with T the kinetic part, which by definition vanishes for constant φ . Clearly, the key quantity $\dot{\varphi}$ must appear in T . To have a chance of implementing **Poincaré invariance**, the whole vector $\partial_\mu\varphi$ must then appear. Moreover, focussing again on small excitations, we are most interested in the lowest terms in an expansion of T in $\partial_\mu\varphi$.

It is a fairly obvious fact that $(\partial_\mu\varphi)(\partial_\nu\varphi)\eta^{\mu\nu}$ is the lowest-order Poincaré-invariant expression suitable for \mathcal{L}_{kin} . Thus, the unique (up to rescaling of φ), lowest-order action reads

$$S = \int d^4x \mathcal{L} = \int d^4x \left(\frac{1}{2}(\partial_\mu\varphi)(\partial^\mu\varphi) - \frac{m^2}{2}\varphi^2 \right). \quad (2.10)$$

It is also a useful approximation to many more interesting systems.

To gain some intuition, separate time and space ($\varphi(x) = \varphi(t, \vec{x})$) and discretise the latter:

$$\int d^3x \longrightarrow \sum_{\vec{x}} \quad (2.11)$$

where $\vec{x} \in (3\text{-dimensional, cubic lattice with spacing } \Delta)$. The Lagrangian

$$L = \int d^3x \left(\frac{1}{2}\dot{\varphi}^2 - \frac{1}{2}(\nabla\varphi)^2 - \frac{m^2}{2}\varphi^2 \right) \quad (2.12)$$

then turns into³

$$L = \underbrace{\sum_{\vec{x}} \left\{ \frac{1}{2}\dot{\varphi}(t, \vec{x})^2 \right\}}_{T_{\text{class. mech.}}} - \underbrace{\frac{1}{2} \sum_{i=1}^3 \left(\frac{\varphi(t, \vec{x} + \hat{e}_i\Delta) - \varphi(t, \vec{x})}{\Delta} \right)^2 - \frac{m^2}{2}\varphi^2}_{V_{\text{class. mech.}}}. \quad (2.13)$$

³ We are deliberately dropping a factor Δ^3 which would normally come with the sum, and which then also affects the normalisation of the canonical momentum associated with φ . The reader is invited to do this more carefully – for our purposes this naive discretisation is sufficient.

Note the difference between the field-theoretic kinetic term T defined above and the term $T_{\text{class. mech.}}$ which corresponds to the kinetic term of the analogous mechanical system. This mechanical system has infinitely many degrees of freedom, but this can be cured by considering a finite-volume box. Its dynamics is (due to our restriction to quadratic order in φ) simply that of a set of coupled harmonic oscillators. It will be easy to decouple them - see below.

Returning to the continuum, the equation of motion is easily derived as follows

$$\begin{aligned} 0 = \delta S &= \delta \int d^4x \left(\frac{1}{2}(\partial\varphi)^2 - \frac{m^2}{2}\varphi^2 \right) \\ &= \int d^4x \left((\partial_\mu\varphi)\eta^{\mu\nu}(\partial_\nu\delta\varphi) - m^2\varphi\delta\varphi \right) \\ &= \int d^4x \left(-(\partial_\nu\partial_\mu\varphi)\eta^{\mu\nu} - m^2\varphi \right) \delta\varphi. \end{aligned} \quad (2.14)$$

Hence, we obtain the **Klein-Gordon-equation**

$$(\partial^2 + m^2)\varphi = 0, \quad (2.15)$$

where $\partial^2 \equiv \partial_\mu\partial^\mu$. Its solutions are plane waves, e.g.

$$\varphi(x) = \varphi_0 \sin(kx) \quad \text{with} \quad k^2 - m^2 = 0 \quad (2.16)$$

Here $k = (k^0, \vec{k})$ characterises the 4-momentum of the corresponding particle or particles. Choosing $k = (k^0, \vec{0})$ implies that the particles are at rest. The quantity $k^0 = m$ will turn out to be the particle mass.

2.2 Classical theory - Hamiltonian formulation

In classical mechanics, the transition from Lagrange to Hamilton is accomplished according to

$$L(q_i, \dot{q}_i) \quad \longrightarrow \quad H(q_i, p_i) = \sum_i p_i \dot{q}_i - L \quad \text{with} \quad p_i = \frac{\partial L(q_i, \dot{q}_i)}{\partial \dot{q}_i} \quad (2.17)$$

In field theory on the lattice, as discussed above, we have

$$L(\varphi, \dot{\varphi}) = \frac{1}{2} \sum_{\vec{x}} \dot{\varphi}(\vec{x})^2 + \dots \quad (2.18)$$

Here \vec{x} plays the role of the index i and we suppressed the terms without time-derivatives for brevity. For the conjugate momentum

$$\pi(\vec{x}) \equiv \frac{\partial L}{\partial \dot{\varphi}(\vec{x})} = \dot{\varphi}(\vec{x}). \quad (2.19)$$

This follows simply from

$$\frac{\partial}{\partial \dot{\varphi}(\vec{x})} \left(\frac{1}{2} \sum_{\vec{y}} \dot{\varphi}(\vec{y})^2 \right) = \dot{\varphi}(\vec{x}). \quad (2.20)$$

It is common to call the canonically conjugate field π rather than p since this is a better match for the greek variable φ and also since p is frequently used for the 4-momentum of the particles, to which we will soon come.

The Hamiltonian then reads

$$\begin{aligned} \Rightarrow H &= \sum_{\vec{x}} \pi(\vec{x}) \dot{\varphi}(\vec{x}) - L \\ &= \sum_{\vec{x}} \left\{ \pi^2(\vec{x}) - \frac{1}{2} \pi^2(\vec{x}) + \frac{1}{2} \sum_{i=1}^3 \left(\frac{\varphi(t, \vec{x} + \hat{e}_i \Delta) - \varphi(t, \vec{x})}{\Delta} \right)^2 + \frac{m^2}{2} \varphi^2 \right\} \\ &= \sum_{\vec{x}} \left\{ \frac{1}{2} \pi^2(\vec{x}) + \frac{1}{2} \sum_{i=1}^3 \left(\frac{\varphi(t, \vec{x} + \hat{e}_i \Delta) - \varphi(t, \vec{x})}{\Delta} \right)^2 + \frac{m^2}{2} \varphi^2 \right\}. \end{aligned} \quad (2.21)$$

In continuum field theory, we clearly need to generalize $\partial L / \partial q_i$ to the case of a continuous index i . Thus, we need functional derivatives, which we briefly recall:

Let $F : f \mapsto \mathbb{R}$ be a functional. The functional derivative $\delta F / \delta f(x)$ is defined by:

$$F[f + \epsilon] - F[f] = \int dx \frac{\delta F[f]}{\delta f(x)} \epsilon(x) + O(\epsilon^2). \quad (2.22)$$

We also use the natural generalization

$$\sum_i \pi_i \dot{q}_i \longrightarrow \int d^3x \pi(\vec{x}) \dot{\varphi}(\vec{x}) \quad (2.23)$$

The transition from Lagrange to Hamiltonian for continuous systems then reads

$$\pi(\vec{x}) = \frac{\delta}{\delta \dot{\varphi}(\vec{x})} L[\varphi, \dot{\varphi}] \quad , \quad H[\varphi, \pi] = \int d^3x \pi \dot{\varphi} - L. \quad (2.24)$$

From our definition of functional derivatives above, one easily shows that for a functional of the form

$$F = \int dx A(f(x)) \quad (2.25)$$

one has

$$\frac{\delta F}{\delta f(x)} = A'(f(x)). \quad (2.26)$$

Hence

$$\pi(\vec{x}) = \frac{\partial \mathcal{L}}{\partial \dot{\varphi}}(\vec{x}) = \dot{\varphi}(\vec{x}) \quad (2.27)$$

and

$$H[\varphi, \pi] = \frac{1}{2} \int d^3x \underbrace{\left(\pi^2 + (\vec{\nabla} \varphi)^2 + m^2 \varphi^2 \right)}_{2\mathcal{H}} = \int d^3x \mathcal{H}, \quad (2.28)$$

with \mathcal{H} the **Hamiltonian density**.

2.3 Quantisation: real scalar

In our conventions with $\hbar = 1$, canonical quantisation proceeds by promoting observables to operators and postulating

$$[q_i, p_j] = i\delta_{ij} \quad , \quad [q_i, q_j] = [p_i, p_j] = 0. \quad (2.29)$$

We could apply this to the lattice version of our model, but there is no complication in moving on directly to the continuum case, where the analogous relations read

$$[\varphi(\vec{x}), \pi(\vec{y})] = i\delta^3(\vec{x} - \vec{y}) \quad , \quad [\varphi(\vec{x}), \varphi(\vec{y})] = [\pi(\vec{x}), \pi(\vec{y})] = 0. \quad (2.30)$$

With H given as in (2.28), the similarity to a set of harmonic oscillators is obvious. The different oscillators are coupled to each other by the gradient term. The reason is that $\vec{\nabla}\varphi(t, \vec{x})$ involves, through its defining limiting procedure, fields at different points. The oscillators may be decoupled by a Fourier transformation,

$$\varphi(\vec{x}) = \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\vec{x}} \tilde{\varphi}(\vec{p}) \quad , \quad \tilde{\varphi}(\vec{p}) = \int d^3x e^{-i\vec{p}\vec{x}} \varphi(\vec{x}), \quad (2.31)$$

with analogous relations between $\pi(\vec{x})$ and $\tilde{\pi}(\vec{p})$. The commutation relations for $\tilde{\varphi}$, $\tilde{\pi}$ read:

$$\begin{aligned} [\tilde{\varphi}(\vec{p}), \tilde{\pi}(\vec{q})] &= \int d^3x d^3y e^{-i\vec{p}\vec{x}} e^{-i\vec{q}\vec{y}} [\varphi(\vec{x}), \pi(\vec{y})] = i \int d^3x e^{-i(\vec{p}+\vec{q})\vec{x}} = i(2\pi)^3 \delta^3(\vec{p} + \vec{q}) \\ [\tilde{\varphi}(\vec{p}), \tilde{\varphi}(\vec{q})] &= [\tilde{\pi}(\vec{p}), \tilde{\pi}(\vec{q})] = 0. \end{aligned} \quad (2.32)$$

Now, we express H through $\tilde{\varphi}$, $\tilde{\pi}$. We focus on the most interesting term, $(\vec{\nabla}\varphi)^2$:

$$\int d^3x (\vec{\nabla}\varphi)^2 = \int d^3x \int \frac{d^3p}{(2\pi)^3} i\vec{p}_k e^{i\vec{p}\vec{x}} \tilde{\varphi}(\vec{p}) \int \frac{d^3q}{(2\pi)^3} i\vec{q}_k e^{i\vec{q}\vec{x}} \tilde{\varphi}(\vec{q}). \quad (2.33)$$

The x -integration gives a δ -function, $(2\pi)^3 \delta^3(\vec{p} + \vec{q})$, allowing us to trivially perform the q -integration and to obtain

$$\int \frac{d^3p}{(2\pi)^3} \vec{p}^2 \tilde{\varphi}(\vec{p}) \tilde{\varphi}(-\vec{p}). \quad (2.34)$$

We note that the reality of our field, $\overline{\varphi(\vec{x})} = \varphi(\vec{x})$, implies the relation $\overline{\tilde{\varphi}(\vec{p})} = \tilde{\varphi}(-\vec{p})$ for its Fourier transform. (Jumping ahead, we also note that an analogous logic continues to hold for the operators: $\varphi(\vec{x})^\dagger = \varphi(\vec{x})$ implies $\tilde{\varphi}(\vec{p})^\dagger = \tilde{\varphi}(-\vec{p})$.) With this, and after similar but simpler manipulations for the other terms of (2.28), we eventually find

$$H = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2} (|\tilde{\pi}|^2 + (\vec{p}^2 + m^2)|\tilde{\varphi}|^2). \quad (2.35)$$

Here $|\tilde{\pi}|^2$ can be understood either literally or as $\tilde{\pi}\tilde{\pi}^\dagger$, depending on whether we are before or after quantisation.

To further perfect the similarity to harmonic oscillators, let us introduce

$$\omega_{\vec{p}} \equiv \sqrt{\vec{p}^2 + m^2}, \quad (2.36)$$

such that

$$H = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2} (|\tilde{\pi}|^2 + \omega_{\vec{p}}^2 |\tilde{\varphi}|^2) . \quad (2.37)$$

At this point, it may be useful to collect the relevant equations from quantum mechanics. Using π for the canonical momentum to make the analogy to field theory more obvious, we have the Hamiltonian and the commutation relations

$$H = \frac{1}{2} (\pi^2 + \omega^2 q^2) \quad , \quad [q, \pi] = i , \quad (2.38)$$

the definition of creation and annihilation operators

$$a = \frac{1}{2} \left(\sqrt{2\omega} q + i \sqrt{\frac{2}{\omega}} \pi \right) \quad , \quad a^\dagger = \frac{1}{2} \left(\sqrt{2\omega} q - i \sqrt{\frac{2}{\omega}} \pi \right) , \quad (2.39)$$

and finally the resulting new expression for the Hamiltonian as well as the new commutation relations

$$H = \omega \left(a^\dagger a + \frac{1}{2} \right) \quad , \quad [a, a^\dagger] = 1 . \quad (2.40)$$

Motivated by this, we may hope that (2.37) will simplify similarly to quantum mechanics if we define

$$a_{\vec{p}} = \frac{1}{2} \left(\sqrt{2\omega_{\vec{p}}} \tilde{\varphi}(\vec{p}) + i \sqrt{\frac{2}{\omega_{\vec{p}}}} \tilde{\pi}(\vec{p}) \right) . \quad (2.41)$$

$$a_{\vec{p}}^\dagger = \frac{1}{2} \left(\sqrt{2\omega_{\vec{p}}} \tilde{\varphi}(-\vec{p}) - i \sqrt{\frac{2}{\omega_{\vec{p}}}} \tilde{\pi}(-\vec{p}) \right) . \quad (2.42)$$

As a side remark, please note that it is not completely obvious a priori that such a simplification will occur. The reason is that our analogy to the oscillator is not perfect: Unlike p, q our $\tilde{\varphi}, \tilde{\pi}$ are not real and they are not (quite) conjugate variables: $\delta^3(\vec{p} + \vec{q}) = \delta^3(\vec{p} - (-\vec{q}))$, i.e. $\tilde{\varphi}(\vec{p})$ is conjugate to $\tilde{\pi}(-\vec{p})$. We could have kept massaging our system into perfect agreement with a set of oscillators before introduce a, a^\dagger , but we would not gain much new information. The ansatz above is natural enough and it is simpler to directly check that it works.

Indeed, it is easy to derive that

$$\left[a_{\vec{p}}, a_{\vec{q}}^\dagger \right] = (2\pi)^3 \delta^3(\vec{p} - \vec{q}) \quad , \quad [a_{\vec{p}}, a_{\vec{q}}] = [a_{\vec{p}}^\dagger, a_{\vec{q}}^\dagger] = 0 . \quad (2.43)$$

Furthermore, one finds

$$\tilde{\varphi}(\vec{p}) = \frac{1}{\sqrt{2\omega_{\vec{p}}}} (a_{\vec{p}} + a_{-\vec{p}}^\dagger) \quad , \quad \tilde{\pi}(\vec{p}) = -i \sqrt{\frac{\omega_{\vec{p}}}{2}} (a_{\vec{p}} - a_{-\vec{p}}^\dagger) \quad (2.44)$$

and hence:

$$H = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2} \left(\frac{\omega_{\vec{p}}}{2} (a_{\vec{p}} - a_{-\vec{p}}^\dagger)(a_{\vec{p}}^\dagger - a_{-\vec{p}}) + \frac{\omega_{\vec{p}}^2}{2\omega_{\vec{p}}} (a_{\vec{p}} + a_{-\vec{p}}^\dagger)(a_{\vec{p}}^\dagger + a_{-\vec{p}}) \right) . \quad (2.45)$$

Cross-terms like $a_{\vec{p}} a_{-\vec{p}}$ cancel. Freely using the substitution $\vec{p} \rightarrow -\vec{p}$, we derive

$$\begin{aligned} H &= \int \frac{d^3p}{(2\pi)^3} \frac{\omega_{\vec{p}}}{2} \left(a_{\vec{p}}^\dagger a_{\vec{p}} + a_{-\vec{p}} a_{-\vec{p}}^\dagger \right) \\ &= \int \frac{d^3p}{(2\pi)^3} \omega_{\vec{p}} \left(a_{\vec{p}}^\dagger a_{\vec{p}} + \frac{1}{2} [a_{\vec{p}}, a_{\vec{p}}^\dagger] \right). \end{aligned} \quad (2.46)$$

The second term may be manipulated, in a slightly hand-wavy way, as follows:

$$(2\pi)^3 \delta^3(\vec{0}) \stackrel{‘=’}{=} \int d^3x e^{i\vec{0}\cdot\vec{x}} \stackrel{‘=’}{=} \text{Vol}(\mathbb{R}^3) \equiv V. \quad (2.47)$$

One then obtains

$$H = \int \frac{d^3p}{(2\pi)^3} \omega_{\vec{p}} a_{\vec{p}}^\dagger a_{\vec{p}} + V \int \frac{d^3p}{(2\pi)^3} \frac{1}{2} \omega_{\vec{p}}. \quad (2.48)$$

The volume-factor V can be derived more properly by using T^3 instead of \mathbb{R}^3 , where momenta become discrete and $\delta^3(\vec{p} - \vec{q})$ becomes $\delta_{\vec{p},\vec{q}}$. The integral multiplying V is truly divergent and requires the introduction of a so-called **UV-cutoff**. This **UV-divergence** is due to contributions of zero-point energies of harmonic oscillators with arbitrarily high frequencies to the total vacuum energy density.

Of course, the **vacuum energy** represented by the last term in (2.48) is irrelevant in the pure QFT context we are in at the moment. It can be absorbed in an overall constant shift of H . However, if our QFT lives not in a fixed \mathbb{R}^3 but in a dynamical space, i.e. if our QFT is coupled to gravity, then this vacuum energy can curve space-time, inducing a **cosmological constant**. This is one of the aspects of the so-called **cosmological constant problem**.

In non-trivial geometries (e.g. QED with conducting plates) the energies of the low-lying modes can be manipulated by moving the plates. This leads to a finite effect (force on the plates) independently of the still present divergence of $|\vec{p}| \rightarrow \infty$. This is known as the **Casimir energy/effect**, another situation where the vacuum energy is manifestly physical. More details can be found in the problems and e.g. in [2, 3].

We now leave the issue of the vacuum energy and return to our main line of development. We may summarise what we have found by the relations

$$\varphi(\vec{x}) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{p}}}} e^{i\vec{p}\vec{x}} (a_{\vec{p}} + a_{-\vec{p}}^\dagger) \quad (2.49)$$

$$\pi(\vec{x}) = -i \int \frac{d^3p}{(2\pi)^3} \sqrt{\frac{\omega_{\vec{p}}}{2}} e^{i\vec{p}\vec{x}} (a_{\vec{p}} - a_{-\vec{p}}^\dagger) \quad (2.50)$$

with

$$[a_{\vec{p}}, a_{\vec{q}}^\dagger] = (2\pi)^3 \delta^3(\vec{p} - \vec{q}) \quad \text{and} \quad H = \int \frac{d^3p}{(2\pi)^3} \omega_{\vec{p}} a_{\vec{p}}^\dagger a_{\vec{p}}. \quad (2.51)$$

At the moment, this is just an (operator) algebra with one distinguished operator H . Physics starts if we also provide a Hilbert-space representation of this algebra. To construct this representation, we postulate a **vacuum state** $|0\rangle$ such that

$$a_{\vec{p}} |0\rangle = 0 \quad \forall \vec{p}. \quad (2.52)$$

So-called **one-particle states** are defined by $a_{\vec{p}}^\dagger|0\rangle$ (for any \vec{p}). It is easy to calculate their energy:

$$H a_{\vec{p}}^\dagger|0\rangle = \int \frac{d^3k}{(2\pi)^3} \omega_{\vec{k}} a_{\vec{k}}^\dagger a_{\vec{k}} a_{\vec{p}}^\dagger|0\rangle = \int \frac{d^3k}{(2\pi)^3} \omega_{\vec{k}} a_{\vec{k}}^\dagger (2\pi)^3 \delta^{(3)}(\vec{k} - \vec{p})|0\rangle = \omega_{\vec{p}} a_{\vec{p}}^\dagger|0\rangle . \quad (2.53)$$

We emphasise that this is an important result:

$$H a_{\vec{p}}^\dagger|0\rangle = \omega_{\vec{p}} a_{\vec{p}}^\dagger|0\rangle . \quad (2.54)$$

Indeed, $\omega_{\vec{p}} = \sqrt{\vec{p}^2 + m^2}$ is the special-relativistic energy of a particle with mass m and momentum \vec{p} , justifying our name ‘one particle state’.

Two-particle states are defined by $a_{\vec{p}}^\dagger a_{\vec{q}}^\dagger|0\rangle$ (for any \vec{p} and any \vec{q}). It is easy to check that

$$H a_{\vec{p}}^\dagger a_{\vec{q}}^\dagger|0\rangle = (\omega_{\vec{p}} + \omega_{\vec{q}}) a_{\vec{p}}^\dagger a_{\vec{q}}^\dagger|0\rangle . \quad (2.55)$$

The calculation proceeds by commuting the annihilation operator $a_{\vec{k}}$ from H to the right until it hits the vacuum. Each time it passes one of the creation operators, we pick up a δ^3 distribution, and hence an $\omega_{\vec{k}}$. The term in brackets is the energy of two non-interacting particles.

This extends to any number of particles and the total space spanned by all those states is called the **Fock space**.

Let us normalise the vacuum:

$$\left| |0\rangle \right|^2 = \langle 0|0\rangle = 1 . \quad (2.56)$$

One then finds

$$\left(a_{\vec{p}}^\dagger|0\rangle \right) \cdot \left(a_{\vec{q}}^\dagger|0\rangle \right) = \langle 0| a_{\vec{p}} a_{\vec{q}}^\dagger|0\rangle = (2\pi)^3 \delta^{(3)}(\vec{p} - \vec{q}) , \quad (2.57)$$

i.e. one-particle states with different momenta are orthogonal, with so-called δ -function normalisation. Analogous formulae for two- and more particle states are easy to derive. States with different numbers of particles are always orthogonal, independently of the momenta.

A convenient notation and normalisation convention is the following:

$$\begin{aligned} |\vec{p}\rangle &= \sqrt{2\omega_{\vec{p}}} a_{\vec{p}}^\dagger|0\rangle \\ |\vec{p}\vec{q}\rangle &= \sqrt{2\omega_{\vec{p}}} \sqrt{2\omega_{\vec{q}}} a_{\vec{p}}^\dagger a_{\vec{q}}^\dagger|0\rangle \\ &\dots \end{aligned} \quad (2.58)$$

It implies that

$$\langle \vec{p}|\vec{q}\rangle = 2\omega_{\vec{p}} (2\pi)^3 \delta^{(3)}(\vec{p} - \vec{q}) . \quad (2.59)$$

Note that many books absorb the factor $\sqrt{2\omega_{\vec{p}}}$ into the definition of the $a_{\vec{p}}$, such that $\sqrt{2\omega_{\vec{p}}} a_{\vec{p}} \rightarrow a_{\vec{p}}$. We will also do so after we understood the issue of Lorentz covariance, to which this is related, better. At the moment, our present convention for the creation and annihilation operators appears more natural since it follows directly from the quantum-mechanical harmonic oscillator.

2.4 Quantisation: complex scalar

Our real-scalar lagrangian immediately generalizes to the complex case,

$$\mathcal{L} = \eta^{\mu\nu}(\partial_\mu\phi)(\partial_\nu\bar{\phi}) - m^2\phi\bar{\phi} = |\partial\phi|^2 - m^2|\phi|^2, \quad (2.60)$$

where the second notation is slightly sloppy but very common. Note the different normalisation conventions (the missing factor 1/2) compared to the real case.

With $\phi = (\varphi_1 + i\varphi_2)/\sqrt{2}$ we recover precisely twice the real-scalar lagrangian analysed before. Thus, we could appeal to the last section and declare our new Fock space simply to be the tensor product of two Fock spaces of the previous type. The Hamiltonian would be the sum of two copies of the previous Hamiltonian, and we would be done without any new calculations.

Nevertheless, it is also useful to quantise this system without giving up the complex notation. Before doing so, let us go through the classical analysis. Our previous derivation of the equations of motion generalizes, in the case of several independent fields φ_i , to

$$\delta S = 0 \quad \Rightarrow \quad \frac{\partial\mathcal{L}}{\partial\varphi_i} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_i)} = 0, \quad \forall i. \quad (2.61)$$

Treating $\phi, \bar{\phi}$ as independent variables, this implies

$$-m^2\bar{\phi} - \square\bar{\phi} = 0, \quad -m^2\phi - \square\phi = 0 \quad (2.62)$$

where $\square = \partial_\mu\partial^\mu$ and the second equation either follows from the first by complex conjugation or by variation of S w.r.t. $\bar{\phi}$. The conjugate momenta take the form

$$\pi = \frac{\partial\mathcal{L}}{\partial\dot{\phi}} = \dot{\bar{\phi}} \quad \bar{\pi} = \frac{\partial\mathcal{L}}{\partial\dot{\bar{\phi}}} = \dot{\phi}. \quad (2.63)$$

We thus get

$$\mathcal{H} = \pi\dot{\phi} + \bar{\pi}\dot{\bar{\phi}} - \mathcal{L} = |\pi|^2 + \left|\vec{\nabla}\phi\right|^2 + m^2|\phi|^2. \quad (2.64)$$

Quantisation proceeds in the standard way, by imposing that

$$[\phi(\vec{x}), \pi(\vec{y})] = [\phi^\dagger(\vec{x}), \pi^\dagger(\vec{y})] = i\delta^{(3)}(\vec{x} - \vec{y}) \quad (2.65)$$

and all other commutators vanish. Before, we were successful with the ansatz

$$\varphi(\vec{x}) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{p}}}} e^{i\vec{p}\vec{x}} (a_{\vec{p}} + a_{-\vec{p}}^\dagger) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{p}}}} \left(a_{\vec{p}}^\dagger e^{-i\vec{p}\vec{x}} + a_{\vec{p}} e^{i\vec{p}\vec{x}} \right). \quad (2.66)$$

In this ansatz the reality of $\varphi(\vec{x})$ is encoded in $(a_{\vec{p}}^\dagger) = (a_{\vec{p}})^\dagger$. This suggest the new ansatz

$$\phi(\vec{x}) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{p}}}} \left(a_{\vec{p}}^\dagger e^{-i\vec{p}\vec{x}} + b_{\vec{p}} e^{i\vec{p}\vec{x}} \right), \quad (2.67)$$

where $b_{\vec{p}}$ is a set of independent operators. With this ansatz one finds commutation relations for $\phi, \pi, \phi^\dagger, \pi^\dagger$, that are the commutation relations of *two* sets of harmonic

oscillators described by $a_{\vec{p}}^\dagger, a_{\vec{p}}$ and $b_{\vec{p}}^\dagger, b_{\vec{p}}$ (see problem sheet). Note that we now write ϕ^\dagger rather than $\bar{\phi}$ since we have moved to the quantum theory. The Hamiltonian reads

$$H = \int \frac{d^3p}{(2\pi)^3} \omega_{\vec{p}} \left(a_{\vec{p}}^\dagger a_{\vec{p}} + b_{\vec{p}}^\dagger b_{\vec{p}} \right) \quad (2.68)$$

and the Fock space is now built with *two* types of particles, created by $a_{\vec{p}}^\dagger$ and by $b_{\vec{p}}^\dagger$. These will turn out to be particles and anti-particles. Also, we immediately see that our system has a symmetry: $\phi \rightarrow e^{i\alpha}\phi$. We will see that this gives us a conserved charge.

3 Noether's theorem

3.1 Formulation and derivation in field theory

With every continuous symmetry of the action comes a **conserved current density** (and a conserved charge). This is very similar to Noether's theorem in mechanics. The crucial novelty is the current.

To derive the theorem, we assume that, infinitesimally, our continuous symmetry transformation takes the form

$$\varphi(x) \rightarrow \varphi'(x) = \varphi(x) + \varepsilon\chi(x). \quad (3.1)$$

We denote this change of φ by $\delta_\varepsilon\varphi \equiv \varphi' - \varphi$. The induced change of \mathcal{L} reads

$$\delta_\varepsilon\mathcal{L} \equiv \mathcal{L}' - \mathcal{L} = \mathcal{L}(\varphi', \partial\varphi') - \mathcal{L}(\varphi, \partial\varphi). \quad (3.2)$$

The fact that our transformation is a *symmetry* implies that \mathcal{L} changes only by a total derivative,

$$\delta_\varepsilon\mathcal{L} = \varepsilon\partial_\mu F^\mu(\varphi, \partial\varphi, \partial\partial\varphi, x), \quad (3.3)$$

where F is some appropriately chosen vector field. Note that, as emphasised in the expression above, F can in general depend on higher derivatives of φ than what appears in the lagrangian, and it can also explicitly depend on x . This can occur even though \mathcal{L} usually does not have an explicit x -dependence.

The statement that (3.3) defines a symmetry is very similar to the familiar situation in classical mechanics. Still, it may be useful to check this claim also in our case: We assume that φ satisfies the equation of motion and check that, under the above conditions, φ' does so too. For this, we need to check that $\delta S' = 0$ for any variation $\delta\varphi$ of φ' in a bounded region. Indeed,

$$\delta S' = \delta S + \delta(\delta_\varepsilon S) = \delta(\delta_\varepsilon S) = \int d^4x \delta(\delta_\varepsilon\mathcal{L}) = \int d^4x \varepsilon\partial_\mu \delta F^\mu(x) = 0 \quad (3.4)$$

where the last equality follows by Gauss law since $\delta\varphi$ and hence δF vanish outside a bounded region. Note that we did not demand that $\delta_\varepsilon S = 0$ or that $\delta_\varepsilon\varphi$ should vanish outside a bounded region.

A simple calculation now leads to the statement of the theorem. First,

$$\begin{aligned}\varepsilon\partial_\mu F^\mu &= \delta_\varepsilon\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\varphi}\delta_\varepsilon\varphi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\delta_\varepsilon\partial_\mu\varphi \\ &= \frac{\partial\mathcal{L}}{\partial\varphi}\delta_\varepsilon\varphi + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\delta_\varepsilon\varphi\right) - \left(\partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\right)\delta_\varepsilon\varphi.\end{aligned}\quad (3.5)$$

Using the equation of motion

$$\frac{\partial\mathcal{L}}{\partial\varphi} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)} = 0 \quad (3.6)$$

and $\delta_\varepsilon\varphi = \varepsilon\chi$ one finds

$$\varepsilon\partial_\mu F^\mu = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\varepsilon\chi\right). \quad (3.7)$$

Thus, the current

$$j^\mu \equiv \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\chi - F^\mu \quad (3.8)$$

is conserved, $\partial_\mu j^\mu = 0$.

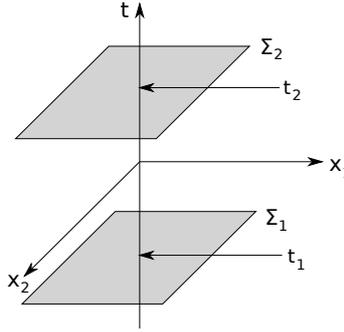


Figure 3: Charge conservation.

For many field configurations the following integral can be defined:

$$Q(t) \equiv \int d^3x j^0(t, \vec{x}). \quad (3.9)$$

We call it the conserved charge since (cf. Fig. 3)

$$Q(t_2) - Q(t_1) = \int_{\Sigma_2} df_\mu j^\mu - \int_{\Sigma_1} df_\mu j^\mu = \int_{\text{Vol}} dV \partial_\mu j^\mu = 0. \quad (3.10)$$

Here the first equality holds because the normal vectors pick out the time-component of j , and the second follows from Gauß' law.

As an exercise, the readers may try to derive $\dot{Q} = 0$ directly, i.e. without introducing a finite interval $\delta t = t_2 - t_1$. Moreover, they may apply our derivation to obtain the familiar Noether theorem of mechanics.

3.2 Energy-Momentum-Conservation

The symmetry transformation underlying energy-momentum-conservation is

$$x^\mu \rightarrow x'^\mu = x^\mu - \epsilon^\mu, \quad (3.11)$$

i.e. a translation by a small amount which we chose as $-\epsilon$. Actually, there are four symmetries, corresponding to the fact that our ϵ is a 4-vector.

From this the transformation behavior of the field and the Lagrangian follows:

$$\begin{aligned} \varphi'(x) = \varphi(x + \epsilon) &\Rightarrow \delta_\epsilon \varphi = \varphi(x + \epsilon) - \varphi(x) \simeq \epsilon^\nu \partial_\nu \varphi \\ \mathcal{L}'(x) = \mathcal{L}(x + \epsilon) &\Rightarrow \delta_\epsilon \mathcal{L} = \mathcal{L}(x + \epsilon) - \mathcal{L}(x) \simeq \epsilon^\mu \partial_\mu \mathcal{L} \\ &= \epsilon^\nu \partial_\mu (\delta^\mu_\nu \mathcal{L}) = \epsilon^\nu \partial_\mu (F^\mu_\nu). \end{aligned} \quad (3.12)$$

Here we have defined

$$F^\mu_\nu \equiv \delta^\mu_\nu \mathcal{L}. \quad (3.13)$$

The last expression in (3.12) should be understood as a linear superposition of four contributions of the type $\epsilon \partial_\mu F^\mu$. Each of the four contributions has a different infinitesimal parameter ϵ and a different current. They are labelled by the $\nu = 0, 1, 2, 3$:

$$\epsilon^\nu \partial_\mu (F^\mu_\nu) = \sum_\nu \epsilon^\nu \partial_\mu (F^\mu_\nu). \quad (3.14)$$

Similarly, the last expression in the first line of (3.12) is to be read as a linear superposition of four contributions of the type $\epsilon \chi(x)$. Thus, we identify the different χ 's as

$$\chi_\nu = \partial_\nu \varphi. \quad (3.15)$$

We can now apply our general formula, getting four conserved currents labelled by ν :

$$j^\mu_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \chi_\nu - F^\mu_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \partial_\nu \varphi - \delta^\mu_\nu \mathcal{L}. \quad (3.16)$$

This set of currents is also known as the **energy-momentum-tensor** and often written as

$$T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \partial^\nu \varphi - \eta^{\mu\nu} \mathcal{L}. \quad (3.17)$$

It is conserved by construction: $\partial_\mu T^{\mu\nu} = 0$.

The name is justified because the conserved charge associated with the current $T^{\mu 0}$ is the energy:

$$\int d^3x T^{00} = \int d^3x \left(\frac{\partial \mathcal{L}}{\partial \dot{\varphi}} \dot{\varphi} - \mathcal{L} \right) = H \equiv P^0. \quad (3.18)$$

As usual in special relativity, the energy is identical to P^0 , the first component of the 4-vector $\{P^\mu\}$ of the energy-momentum of our field configuration. Similarly, the other conserved charges are the momenta, such that together

$$P^\nu = \int d^3x T^{0\nu}. \quad (3.19)$$

We emphasise that P is a 4-vector in spite of its apparently non-covariant definition. This can be understood if one observes that (cf. our general discussion of the conserved charge) the vector P^ν does not change if the space-like hyperplane used in its definition is rotated. Hence, the definition can be made covariant:

$$P^\nu = \int d^3x T^{0\nu} = \int_\Sigma df_\mu T^{\mu\nu} = \int_{\Sigma'} df_\mu T^{\mu\nu}. \quad (3.20)$$

The spatial components of P explicitly read

$$P^i = \int d^3x T^{0i} = \int d^3x \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} \partial^i \varphi = - \int d^3x \pi \nabla_i \varphi. \quad (3.21)$$

Our expressions are valid classically as well as after quantisation. We may rewrite π and φ in terms a and a^* and then switch to \hat{a} and \hat{a}^\dagger . Here we use the hat-symbol to emphasise the operator-nature. One explicitly finds

$$\hat{P}^i = \int \frac{d^3q}{(2\pi)^3} q^i \hat{a}_q^\dagger \hat{a}_q \quad \text{as well as} \quad \hat{P}^\mu |p\rangle = p^\mu |p\rangle. \quad (3.22)$$

We note that $\hat{a}_q^\dagger \hat{a}_q$ is the **particle number operator**, which makes the above expression for the momenta very intuitive.

The same analysis applied to the complex scalar gives

$$\hat{P}^\mu = \int \frac{d^3q}{(2\pi)^3} q^\mu (\hat{a}_q^\dagger \hat{a}_q + \hat{b}_q^\dagger \hat{b}_q). \quad (3.23)$$

Here we also used that

$$q^0 = \sqrt{\vec{q}^2 + m^2} = \omega_{\vec{q}}. \quad (3.24)$$

With this, our particle-interpretation of the Fock-space is fully justified: We see that \hat{a}_q^\dagger (or in the complex case \hat{a}_q^\dagger and \hat{b}_q^\dagger) create particles/antiparticles with appropriate 4-momentum in the sense of Noether's theorem and of special relativity.

Let us end with a comments concerning the energy-momentum-tensor: From its definition, our $T^{\mu\nu}$ (also known as the 'canonical' energy-momentum tensor) does not necessarily have to be symmetric. It happens to be symmetric for the scalar field, where

$$T^{\mu\nu} = \partial^\mu \varphi \partial^\nu \varphi - \eta^{\mu\nu} \mathcal{L}. \quad (3.25)$$

But this fails already for quantum electrodynamics (QED). However, $T^{\mu\nu}$ can always be *made* symmetric by adding an independently conserved current, which also does not modify the resulting P^ν .

Moreover, in general relativity, one uses the definition

$$T^{\mu\nu}(x) = \frac{-2}{\sqrt{-\det(g)}} \frac{\delta S}{\delta g_{\mu\nu}(x)}, \quad (3.26)$$

where $\det(g)$ stands for the determinant of the matrix $g_{\rho\sigma}$. The action S is formulated using a general metric $g_{\mu\nu}$ instead of the Minkowski metric $\eta_{\mu\nu} = \text{diag}(1, -1, -1, -1)$. This directly gives the symmetric form of $T^{\mu\nu}$. The energy-momentum tensor plays a crucial role in general relativity since

$$S[g] = S[\eta] - \frac{1}{2} \int d^4x h_{\mu\nu} T^{\mu\nu} + \dots \quad (\text{with } g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}) \quad (3.27)$$

characterises the coupling of a QFT to gravity.

3.3 U(1)-symmetry and charge of a complex scalar

Let us consider the already familiar Lagrangian

$$\mathcal{L} = |\partial_\mu \phi|^2 - m^2 |\phi|^2. \quad (3.28)$$

It is invariant under a U(1)-symmetry group of phase rotations, with infinitesimal version

$$\begin{aligned} \phi &\rightarrow \phi' = e^{i\epsilon} \phi = \phi + i\epsilon \phi + \dots \\ \bar{\phi} &\rightarrow \bar{\phi}' = e^{-i\epsilon} \bar{\phi} = \bar{\phi} - i\epsilon \bar{\phi} + \dots \end{aligned} \quad (3.29)$$

As before, we treat $\phi, \bar{\phi}$ as independent fields during the calculation. A rough justification of this convenient and mathematically correct approach is as follows: Consider a Lagrangian with two fields rather than just one field:

$$\mathcal{L} = \eta^{\mu\nu} (\partial_\mu \phi)(\partial_\nu \psi) + \dots \quad \phi, \psi \in \mathbb{C}. \quad (3.30)$$

Do all manipulations in this setting and impose the constraint $\psi = \bar{\phi}$ (the ‘projection on the real subspace’) only at the very end.

In the one-field case, our formula for the Noether current was

$$j^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \chi - F^\mu. \quad (3.31)$$

This has the obvious multi-field generalization

$$j^\mu = \sum_i \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi^i)} \chi^i - F^\mu, \quad (3.32)$$

where i labels the different fields.

In our case, $\mathcal{L} = \mathcal{L}'$ such that $F^\mu = 0$. Moreover, $\{\varphi^i\} = \{\phi, \bar{\phi}\}$. Then our multi-field formula gives

$$\begin{aligned} j^\mu &= \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \chi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \bar{\phi})} \bar{\chi} \quad \text{with } \chi = i\phi, \quad \bar{\chi} = -i\bar{\phi} \\ &= (\partial^\mu \bar{\phi}) i\phi + (\partial^\mu \phi)(-i)\bar{\phi} = -i(\bar{\phi} \overleftrightarrow{\partial}^\mu \phi). \end{aligned} \quad (3.33)$$

Here, in the last expression, we used the common shorthand notation $A \overleftrightarrow{\partial} B \equiv A \partial B - (\partial A) B$. The conserved charge is then given by

$$Q = \int d^3x j^0 = -i \int d^3x \phi^\dagger \overleftrightarrow{\partial}_0 \phi = \int \frac{d^3p}{(2\pi)^3} (a_p^\dagger a_{\bar{p}} - b_p^\dagger b_{\bar{p}}). \quad (3.34)$$

We see that we can think of the states created by a^\dagger/b^\dagger as particles/antiparticles, having the same mass but opposite charge.

4 Heisenberg picture, causality, covariance

4.1 Heisenberg picture

So far, we have worked in the Schrödinger picture and the field operators $\hat{\varphi}(\vec{x}), \hat{\pi}(\vec{x})$. In this picture they have no time dependence. Instead, the dynamics is encoded in our states $|0\rangle, \sqrt{2\omega_{\vec{p}}} a_{\vec{p}}^\dagger |0\rangle = |p\rangle$, etc. These evolve in time in the standard fashion:

$$|p\rangle \equiv |p_{t=0}\rangle \equiv |p_0\rangle \quad \longrightarrow \quad |p_t\rangle = \exp(-iHt) |p_0\rangle . \quad (4.1)$$

We see that the relations between states and field operators which we used so far are, in fact, relations at $t = 0$.

This is clearly not natural for a Poincaré-invariant theory since the time dependence sits in the states while the \vec{x} -dependence sits in the observables. To remedy this, let us convert from the

Schrödinger-picture: O fix and $|\psi_t\rangle = e^{-iHt} |\psi_0\rangle$

to the

Heisenberg-picture: O_t evolves and $|\psi\rangle$ fix.

Here O denotes a generic operator. Its time dependence in the Heisenberg picture follows from the physical requirement

$$\langle \psi_t | O | \psi_t \rangle = \langle \psi | O_t | \psi \rangle . \quad (4.2)$$

This implies

$$O_t = e^{iHt} O e^{-iHt} . \quad (4.3)$$

For us, the most interesting operator is:

$$\varphi(x) = \varphi(t, \vec{x}) = e^{iHt} \left(\int \frac{d^3p}{(2\pi)^3 \sqrt{2\omega_{\vec{p}}}} \left(a_{\vec{p}} e^{i\vec{p}\vec{x}} + a_{\vec{p}}^\dagger e^{-i\vec{p}\vec{x}} \right) \right) e^{-iHt} . \quad (4.4)$$

This simplifies if we commute e^{iHt} through a, a^\dagger , using

$$H a_{\vec{p}} = a_{\vec{p}} (H - \omega_{\vec{p}}) \quad \text{and} \quad H a_{\vec{p}}^\dagger = a_{\vec{p}}^\dagger (H + \omega_{\vec{p}}) . \quad (4.5)$$

To derive the second relation, consider $|\psi\rangle$ such that $H |\psi\rangle = E |\psi\rangle$. From our understanding of the Fock space, we see immediately that

$$H a_{\vec{p}}^\dagger |\psi\rangle = (E + \omega_{\vec{p}}) a_{\vec{p}}^\dagger |\psi\rangle = a_{\vec{p}}^\dagger (H + \omega_{\vec{p}}) |\psi\rangle . \quad (4.6)$$

The first relation follows from hermitian conjugation.

Next, since an exponential of an operator is defined as its Taylor series, we have

$$e^{iHt} a_{\vec{p}} = a_{\vec{p}} e^{i(H - \omega_{\vec{p}})t} \quad \text{and} \quad e^{iHt} a_{\vec{p}}^\dagger = a_{\vec{p}}^\dagger e^{i(H + \omega_{\vec{p}})t} . \quad (4.7)$$

If we also use that $p = \{p^0, \vec{p}\} = \{\omega_{\vec{p}}, \vec{p}\}$ and $px = p^0 x^0 - \vec{p} \cdot \vec{x}$, we arrive at the covariant expression (we will understand the covariance of the integration measure soon)

$$\varphi(x) = \int \frac{d^3p}{(2\pi)^3 \sqrt{2p^0}} \left(a_{\vec{p}} e^{-ipx} + a_{\vec{p}}^\dagger e^{+ipx} \right) . \quad (4.8)$$

We note that, defining $\pi(x) = \pi(t, \vec{x})$ analogously, one finds

$$\pi(x) = \dot{\varphi}(x) \quad \text{and} \quad (\square + m^2) \varphi(x) = 0, \quad (4.9)$$

i.e. the classical relations between canonically conjugate variables and the classical equation of motion hold for Heisenberg-picture field operators.

4.2 Causality

The key statement is that measurements of φ at x and y do not interfere if x and y are space-like separated, i.e.

$$[\varphi(x), \varphi(y)] = 0 \quad \text{for} \quad (x - y)^2 < 0. \quad (4.10)$$

Before deriving this, we note that equal time commutation relations $[\varphi(\vec{x}), \varphi(\vec{y})] = 0$ (in the Schrödinger picture) do not immediately imply causality. However, if we knew that our theory were Poincare-invariant, then causality would follow.

To see this, note first that our causality statement is invariant under $SO^+(1, 3)$. The latter is defined as $SO(1, 3)$ subject to the constraint $\Lambda_0^0 > 0$. Using $SO^+(1, 3)$ we can transform any space-like vector z ($z^2 < 0$) into any other space-like vector of the same length, $z'^2 = z^2$. In particular, if we know that (4.10) holds for $x - y = (0, \vec{x} - \vec{y})$, then it holds for any $x - y$ with $(x - y)^2 < 0$.

However, we do not yet know that our theory (after quantization) is still Poincare. But deriving causality is straightforward without any such assumptions:

$$\begin{aligned} [\varphi(x), \varphi(y)] &= \int \frac{d^3p}{(2\pi)^3 \sqrt{2p^0}} \int \frac{d^3q}{(2\pi)^3 \sqrt{2q^0}} \left\{ [a_{\vec{p}}, a_{\vec{q}}^\dagger] e^{-ipx+iqy} + [a_{\vec{p}}^\dagger, a_{\vec{q}}] e^{+ipx-iqy} \right\} \\ &= \int \frac{d^3p}{(2\pi)^3 2p^0} e^{-ip(x-y)} - \int \frac{d^3p}{(2\pi)^3 2p^0} e^{+ip(x-y)} \\ &= \int \frac{d^4p}{(2\pi)^3} \delta(p^2 - m^2) \Big|_{p^0 > 0} e^{-ip(x-y)} - \int \frac{d^4p}{(2\pi)^3} \delta(p^2 - m^2) \Big|_{p^0 > 0} e^{+ip(x-y)}. \end{aligned} \quad (4.11)$$

Here we used the relation

$$\delta(x^2 - a^2) = \frac{1}{2|a|} [\delta(x + a) + \delta(x - a)]. \quad (4.12)$$

4.3 Covariance

We recall from the previous subsection that, if $x^0 = y^0$, the claim $[\varphi(x), \varphi(y)] = 0$ immediately follows from canonical quantization. It then holds also for generic $x - y$ with $(x - y)^2 < 0$ by Lorentz/Poincaré covariance of the theory. However, we have broken this symmetry in its manifest form during quantization and have not yet established how it reappears at the quantum level. We now understand that this is an important task.

In quantum mechanics, \hat{H} and \hat{P} generate translations in t and \vec{x} , respectively. In QFT, we aim at a Poincare-invariant formulation and it is natural to combine both transformations as

$$e^{i\hat{P}^\mu \delta_\mu} = e^{i\hat{H} \delta_0} e^{-i\hat{P} \vec{\delta}}, \quad (4.13)$$

acting on states $|\psi\rangle$. Note the different signs in the exponents on the right.⁴

In addition to translations, we have to consider Lorentz-rotations,

$$\Lambda^\mu{}_\nu = (e^{i\varepsilon^{\rho\sigma} M_{\rho\sigma}})^\mu{}_\nu = \delta^\mu{}_\nu + i\varepsilon^{\rho\sigma} (M_{\rho\sigma})^\mu{}_\nu + \dots \quad (4.14)$$

Here we introduced a set of 6 independent matrices $M_{\rho\sigma}$ and 6 infinitesimal parameters $\varepsilon^{\rho\sigma}$, in complete analogy to the 3 generators of $SO(3)$ familiar from mechanics and quantum mechanics. By definition, $M_{\rho\sigma} = -M_{\sigma\rho}$ and $\varepsilon^{\rho\sigma} = -\varepsilon^{\sigma\rho}$, so they are indeed 6 rather than 16. In close analogy to $SO(3)$, the matrix $(M_{\rho\sigma})_{\mu\nu}$ is antisymmetric with ± 1 at the positions $\rho\sigma$ and $\sigma\rho$. We leave the details to the problems.

We recall that we have constructed \hat{P} using Noether's theorem: We identified the conserved current associated with small translations, then the conserved charge, then promoted it to an operator. Clearly, with slightly more work the same can be done for the infinitesimal Lorentz transformations generated by $M_{\rho\sigma}$. One eventually arrives at an explicit formula for operator-valued matrices $\hat{M}_{\rho\sigma}$ in terms of creation and annihilation operators. This is worked out in many textbooks but we will not need the explicit expression at the moment. What we need is the understanding that finite translations and Lorentz-transformations can be realised using operators

$$\hat{T} = e^{i\delta^\mu \hat{P}_\mu} \quad \text{and} \quad \hat{\Lambda} = e^{i\varepsilon^{\rho\sigma} \hat{M}_{\rho\sigma}} \quad (4.15)$$

In quantum mechanics, after having realised our symmetry group by unitary operators acting on the Hilbert space, we just need to check that they commute with \hat{H} and we are done. Here, things are slightly different since, of course, boosts $\hat{\Lambda}$ do not commute with \hat{H} (they change the energy of states). We may, however, require invariance of observables in the following sense:⁵

First, measure the field value at some position $x \in \mathbb{R}^{1,3}$ in a state $|\psi\rangle$

$$\langle \psi | \varphi(x) | \psi \rangle \quad (4.16)$$

Now, 'rotate' the state by Λ and measure the field in the rotated position $x'^\mu = \Lambda^\mu{}_\nu x^\nu$. You should get the same result:

$$\langle \psi | \varphi(x) | \psi \rangle = \langle \psi | \hat{\Lambda}^\dagger \varphi(\Lambda x) \hat{\Lambda} | \psi \rangle \quad (4.17)$$

For this to hold, one must demand that operators transform as

$$\hat{\Lambda} \varphi(x) \hat{\Lambda}^\dagger = \varphi(\Lambda x) \quad (4.18)$$

⁴In quantum mechanics, it is easy to convince oneself that $\exp(-i\vec{P}\vec{\delta})$ shifts the wave function by $\vec{\delta}$. Moreover, $\exp(-i\hat{H}\delta_0)|\psi\rangle_t = |\psi\rangle_{t+\delta_0}$, which corresponds to a shift of the time-dependent wave function $|\psi\rangle_t$ by $-\delta_0$. Thus, $\exp(i\hat{H}\delta_0)$ shifts the wave function by δ_0 . So both exponentials together indeed realize a shift by the 4-vector $(\delta_0, \vec{\delta})$.

⁵One may also save the standard quantum mechanical argument, with minimal extra work: While a generic boost generator, let's call it B , does not commute with H , it satisfies $[iH, B] + \partial B/\partial t = 0$ due to its explicit time dependence. Hence $dB/dt = 0$ in the Heisenberg picture. Moreover, one can use this to show that if some state obeys the Schrödinger equation, then so does the boosted state. This justifies calling the boost a symmetry. (I owe this comment to Antonino Di Piazza.)

This can be checked explicitly by expressing $\hat{M}_{\rho\sigma}$ in terms of a, a^\dagger , but we will not do for reasons of time.

It is instructive to check explicitly the consistency with the group law and to compare with the somewhat different relations for classical fields:

We start with operators and apply a first transformation:

$$\hat{\Lambda}_1 \varphi(x) \hat{\Lambda}_1^\dagger = \varphi(\Lambda_1 x). \quad (4.19)$$

A second transformation gives

$$\hat{\Lambda}_2 \left(\hat{\Lambda}_1 \varphi(x) \hat{\Lambda}_1^\dagger \right) \hat{\Lambda}_2^\dagger = \hat{\Lambda}_2 \left(\varphi(\Lambda_1 x) \right) \hat{\Lambda}_2^\dagger = \varphi(\Lambda_2 \Lambda_1 x). \quad (4.20)$$

We see that, as it should be, the transformation by $\hat{\Lambda}_2 \hat{\Lambda}_1$ as an operator corresponds to a transformation of the argument by $\Lambda_2 \Lambda_1$.

Let us contrast this with classical fields, characterised by functions on $\mathbb{R}^{1,3}$. A first transformation reads

$$(\Lambda_1 \varphi)(x) = \varphi(\Lambda_1^{-1} x). \quad (4.21)$$

Here the symbol Λ_1 in $\Lambda_1 \varphi$ denotes a linear operator on the vector space of functions. (We do not give it a hat to avoid confusion with the quantum case.) A second transformation now gives

$$(\Lambda_2 (\Lambda_1 \varphi))(x) = (\Lambda_1 \varphi)(\Lambda_2^{-1} x) = \varphi(\Lambda_1^{-1} \Lambda_2^{-1} x) = \varphi((\Lambda_2 \Lambda_1)^{-1} x) \quad (4.22)$$

This is again consistent with the group law, in spite of the difference concerning the action on the argument.

Finally we would like to have covariance of our Fock-space basis, in the sense that

$$\hat{\Lambda} |p\rangle = |p'\rangle \quad \text{if} \quad p'^\mu = \Lambda^\mu_\nu p^\nu. \quad (4.23)$$

Due to the normalization ambiguity (our states are only δ -function-normalized and the prefactor of $\delta^3(\vec{p} - \vec{q})$ is, in principle, arbitrary) we cannot be a priori be sure about the prefactor in this relation. One can of course work this out explicitly using a, a^\dagger . We will instead only do a consistency check, demonstrating that our normalization convention fits the transformation rule in (4.23):

We demand that

$$\langle p' | q' \rangle \equiv \langle p | \hat{\Lambda}^\dagger \hat{\Lambda} | q \rangle = \langle p | q \rangle, \quad (4.24)$$

where $p' = \Lambda p$, $q' = \Lambda q$. This is equivalent to

$$2p'_0 (2\pi)^3 \delta^3(\vec{p}' - \vec{q}') = 2p_0 (2\pi)^3 \delta^3(\vec{p} - \vec{q}). \quad (4.25)$$

To demonstrate this, we first observe that both sides are non-zero only at the same point. So we just need to check the normalisation, which we can do e.g. by integrating with an arbitrary smooth function in \vec{p} . We chose the integration measure

$$\int \frac{d^3 p}{2p^0} = \int d^4 p \delta(p^2 - m^2) \Big|_{p_0 > 0}. \quad (4.26)$$

Applied to the l.h. side (and dropping the $(2\pi)^3$ factor on both sides) this gives

$$\begin{aligned}
\int \frac{d^3p}{2p^0} 2p'_0 \delta^3(\vec{p}' - \vec{q}') &= \int d^4p \delta(p^2 - m^2) \Big|_{p_0 > 0} 2p'_0 \delta^3(\vec{p}' - \vec{q}') \\
&= \int d^4p' \delta(p'^2 - m^2) \Big|_{p_0 > 0} 2p'_0 \delta^3(\vec{p}' - \vec{q}') \\
&= \int \frac{d^3p'}{2p'^0} 2p'_0 \delta^3(\vec{p}' - \vec{q}') = 1.
\end{aligned} \tag{4.27}$$

Here in the second step we used that $p^2 = p'^2$ and $\det \Lambda = 1$.

On the r.h. side, one simply has

$$\int \frac{d^3p}{2p^0} 2p_0 \delta^3(\vec{p} - \vec{q}) = 1, \tag{4.28}$$

so we are done. For more details see e.g. the books by Itzykson/Zuber (Sec. 3.1.2) and by Weinberg (Vol. 1, Sec. 2).

5 Perturbation theory: first steps in a naive, leading-order approach

5.1 S-Matrix

Until now, the potential part of our scalar-field lagrangian contained only the mass term:

$$V(\varphi) = \frac{m^2}{2} \varphi^2. \tag{5.1}$$

This restriction may be justified by noting that the mass term is the leading non-trivial term in the Taylor expansion around $\varphi = 0$. (The constant term is unobservable unless we include gravity and the linear term can be removed by a redefinition $\varphi \rightarrow \varphi + \text{const.}$)

The resulting theory is ‘free’, which is clear since our Fock space basis was also an energy-eigenstate basis. As a result, neither particle number nor the momenta of particles were changed by the dynamics. In other words: No particle-scattering or decays.

As a next step, it is natural to ask for higher-order terms in the Taylor expansion of V . To simplify the analysis, we will for the moment impose the discrete symmetry $\varphi \rightarrow -\varphi$. In that case the next term in the Taylor expansion is

$$\frac{\lambda}{4!} \varphi^4. \tag{5.2}$$

Adding this term to the potential, our Lagrangian becomes

$$\mathcal{L} = \frac{1}{2} (\partial\varphi)^2 - \frac{1}{2} m^2 \varphi^2 - \frac{\lambda}{4!} \varphi^4. \tag{5.3}$$

We split Lagrangian and Hamiltonian density as

$$\begin{aligned}\mathcal{L} &= \mathcal{L}_0 + \mathcal{L}_{\text{int}} \quad , & \mathcal{L}_{\text{int}} &= -\frac{\lambda}{4!}\varphi^4 \\ \mathcal{H} &= \mathcal{H}_0 + \mathcal{H}_{\text{int}} \quad , & \mathcal{H}_{\text{int}} &= \frac{\lambda}{4!}\varphi^4 ,\end{aligned}\tag{5.4}$$

where \mathcal{L}_0 and \mathcal{H}_0 are the free Lagrangian and Hamiltonian density. Notice the different signs of the interaction term in Lagrangian and Hamiltonian.

Had we not imposed the \mathbb{Z}_2 symmetry $\varphi \rightarrow -\varphi$, the next term would have been

$$V_{\text{int}} = \frac{\lambda}{3!}\varphi^3 .\tag{5.5}$$

One argument that is often brought forward against such a ‘ φ^3 -theory’ is that the potential is unbound from below for $\varphi \rightarrow -\infty$. However, this is not really a problem as we are only interested in perturbation theory around $\varphi = 0$. For this it is sufficient that this so-called vacuum is long-lived – it does not have to be absolutely stable. The true reason for starting with φ^4 -theory lies in its simplicity.

In the following discussion we are going to use the **interaction picture**. As you recall from quantum mechanics, this picture is derived from the Schrödinger picture analogously to Heisenberg picture. The only difference is that one transfers only the **free part** of the time-evolution from states to operators:

$$\begin{array}{lll} \text{Schrödinger picture} & \rightarrow & \text{Interaction picture} \\ \text{Operator } O & \rightarrow & O_t^I = e^{iH_0 t} O e^{-iH_0 t} \\ \text{State } |\psi_t\rangle = e^{-iHt} |\psi\rangle & \rightarrow & |\psi_t^I\rangle = e^{iH_0 t} e^{-iHt} |\psi\rangle . \end{array}\tag{5.6}$$

$$\tag{5.7}$$

One may interpret the last expression as describing the evolution of the interaction picture state between 0 and t' :

$$|\psi_{t'}^I\rangle = e^{iH_0 t'} e^{-iHt'} |\psi_0^I\rangle .\tag{5.8}$$

It is immediate to conclude what the evolution between t and 0 is:

$$|\psi_0^I\rangle = e^{iHt'} e^{-iH_0 t} |\psi_t^I\rangle .\tag{5.9}$$

From this, one reads off the evolution from an arbitrary time t to a different arbitrary time t' . It is conveniently written as

$$|\psi_{t'}^I\rangle = U(t', t) |\psi_t^I\rangle \quad \text{with} \quad U(t', t) = e^{iH_0 t'} e^{-iH(t'-t)} e^{-iH_0 t} ,\tag{5.10}$$

where $U(t', t)$ is the unitary time evolution operator of the interaction picture. You may be already familiar with this from time-dependent perturbation theory in quantum mechanics.

Let us split the time evolution from t to t' into n small steps $\Delta = (t' - t)/n$. Then $U(t', t)$ becomes

$$U(t', t) = U(t', t' - \Delta) \cdot U(t' - \Delta, t' - 2\Delta) \dots U(t + \Delta, t) .\tag{5.11}$$

Then look at one individual step:

$$\begin{aligned} U(t + \Delta, t) &= e^{iH_0(t+\Delta)} e^{-iH\Delta} e^{-iH_0t} = e^{iH_0t} e^{iH_0\Delta} e^{-iH\Delta} e^{-iH_0t} \\ &\simeq e^{iH_0t} e^{-iH_{\text{int}}\Delta} e^{-iH_0t} = e^{-iH_{\text{int}}^I(t)\Delta} . \end{aligned} \quad (5.12)$$

Here in going from the first to the second line we dropped all commutator terms in the Baker-Campbell-Hausdorff formula as they are of second order in Δ :

$$e^{iH_0\Delta} e^{-iH\Delta} = e^{i(H_0-H)\Delta + \mathcal{O}(\Delta^2)} \approx e^{-iH_{\text{int}}\Delta} . \quad (5.13)$$

In the last expression in (13), we introduced the symbol $H_{\text{int}}^I(t)$ for the interaction Hamiltonian H_{int} transformed to the interaction picture:

$$H_{\text{int}}^I(t) = e^{iH_0t} H_{\text{int}} e^{-iH_0t} . \quad (5.14)$$

Now, combining all the time steps, we have

$$\begin{aligned} U(t', t) &\simeq e^{-iH_{\text{int}}^I(t'-\Delta)\Delta} \cdot e^{-iH_{\text{int}}^I(t'-2\Delta)\Delta} \dots e^{-iH_{\text{int}}^I(t)\Delta} \\ &\simeq T e^{-iH_{\text{int}}^I(t'-\Delta)\Delta} \cdot e^{-iH_{\text{int}}^I(t'-2\Delta)\Delta} \dots e^{-iH_{\text{int}}^I(t)\Delta} \\ &\simeq T e^{-iH_{\text{int}}^I(t'-\Delta)\Delta - iH_{\text{int}}^I(t'-2\Delta)\Delta + \dots - iH_{\text{int}}^I(t)\Delta} \\ &= T \exp \left(-i \int_t^{t'} d\tau H_{\text{int}}^I(\tau) \right) , \end{aligned} \quad (5.15)$$

where the last expression assumes the limit $n \rightarrow \infty$ and our formula hence becomes exact. Crucially, in the second line we introduced the so-called **time-ordering operator** T , which has the general definition

$$T\varphi(t_1)\varphi(t_2) = \begin{cases} \varphi(t_1)\varphi(t_2) & \text{if } t_1 \geq t_2 \\ \varphi(t_2)\varphi(t_1) & \text{if } t_2 > t_1 \end{cases} , \quad (5.16)$$

with obvious extension to more than two operators. In words: Any string of operators with time-arguments to the right of T has to be reordered such that the times increase from right to left. In fact, as should be clear from this definition, T is not an operator – the name "time-ordering symbol" would be more accurate. The symbol T simply gives a more detailed specification concerning the operator expression written to its right. The key role of T in (5.15) was that it allowed us to combine the exponents in the step from 2nd to 3rd line. The reason is that under the T symbol the order in which operators are written is clearly irrelevant.

Explicitly, we now have

$$H_{\text{int}}^I(t) = e^{iH_0t} \overbrace{\int d^3x \frac{\lambda}{4!} (\varphi(\vec{x}))^4}^{H_{\text{int}}} e^{-iH_0t} = \int d^3x \frac{\lambda}{4!} (\varphi^I(x))^4 . \quad (5.17)$$

\mathcal{H}_{int}

Here we first wrote the Hamiltonian as an integral of the Hamilton density and then introduced the interaction-picture field φ^I . Notice that this is precisely the

Heisenberg-picture field of the free theory. We will thus be able to apply some of the results from our discussion of the free theory.

Before calculating the scattering or S-matrix, we want to motivate scattering intuitively. For a pair of particles with four-momenta p_1, p_2 , scattered into a state with momenta p'_1, p'_2 , we consider the situation in Fig. 4. Since the interaction $\int d^3x \varphi^4$ includes terms of type $a_{\vec{p}'_1}^\dagger a_{\vec{p}'_2}^\dagger a_{\vec{p}_1} a_{\vec{p}_2}$ we expect it to induce **2-to-2-scattering** at leading order: 2 particles are annihilated – 2 particles with different momenta are created. With higher orders in λ and hence a higher power of φ , in general, we could have **2-to- n -scattering** ($n > 2$). The very definition of scattering requires the particles to be separated at early and late times (t and t'). Separation requires localization, such that we will need wave packets rather than plane waves.

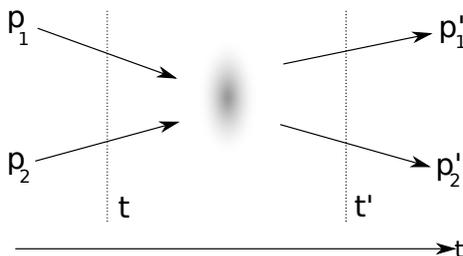


Figure 4: Incoming and outgoing momentum in 2-to-2 scattering. The two particles are separated in the initial and final state.

Now, if the participating particles do not interact at early and late times (as just explained), and since the free single-particle states are time-independent (as we are in the interaction picture), it appears logical to define the scattering matrix as the following limit of the time-evolution operator:

$$S = \lim_{\substack{t \rightarrow -\infty \\ t' \rightarrow \infty}} U(t', t) = T \exp \left(-i \int_{-\infty}^{\infty} dt H_{\text{int}}^I(t) \right) = T \exp \left(i \int d^4x \mathcal{L}_{\text{int}}^I(x) \right). \quad (5.18)$$

One also defines the **S-matrix element**

$$S_{\text{fi}} = \langle p'_1 p'_2 | T \exp \left(i \int d^4x \mathcal{L}_{\text{int}}^I \right) | p_1 p_2 \rangle, \quad (5.19)$$

where f and i denote the final and initial state, respectively. In some problems it is also useful to define the **transition matrix** or **T-matrix**

$$S = \mathbb{1} + iT \quad S_{\text{fi}} = \delta_{\text{fi}} + iT_{\text{fi}}. \quad (5.20)$$

Our S -matrix definition was formally precisely that from quantum mechanics, e.g. for scattering off a localized potential. The required limit exists, roughly speaking, because at early/late times H_{int}^I does not affect the incoming/outgoing states, so the integration range remains finite for the S-matrix-elements between relevant states. In our context, unfortunately, this definition is strictly speaking **wrong**. The reason is that our particles are never ‘far away’ from the scattering potential. Instead, once we look at an interacting theory with $\lambda \neq 0$, also a single, incoming particle is affected by the interaction – the limit above does not make sense. We

could try to give it a meaning by adiabatically switching off λ at early and late times. But one would still have to be very careful about the effect this ‘switching off’ has on an incoming state. Doing all of this properly, which amounts to understanding how to describe single particles in the interacting theory, is known as **LSZ formalism** and will be treated below. For now, we only note that our naive definition of S is completely fine for many ‘leading-order’ applications. And it gives one a lot of easy and correct intuition for how scattering in quantum field theory works. A better definition of S will be supplied later.

Before proceeding, it will be convenient to change the normalization of a, a^\dagger according to

$$(a_{\vec{p}})_{\text{new}} = \sqrt{2\omega_{\vec{p}}}(a_{\vec{p}})_{\text{old}} \quad \text{with} \quad \omega_{\vec{p}} = p^0. \quad (5.21)$$

With these conventions, covariance is emphasized rather than the analogy to harmonic oscillators, which was important to us only at the beginning. With our new a, a^\dagger we have

$$\left[a_{\vec{p}}, a_{\vec{q}}^\dagger \right] = 2p^0(2\pi)^3 \delta^3(\vec{p} - \vec{q}) \quad , \quad |p\rangle = a_{\vec{p}}^\dagger |0\rangle \quad (5.22)$$

and

$$\varphi^I(x) = \int \frac{d^3p}{(2\pi)^3 2p^0} \left(a_{\vec{p}} e^{-ipx} + a_{\vec{p}}^\dagger e^{ipx} \right). \quad (5.23)$$

Please remember to use these new, ‘Lorentz-covariant’ relations in all that follows. Also, recall that the interaction picture fields φ^I are the same as the Heisenberg-picture fields of the free theory.

We note in passing that we do not need to rethink the quantization of our theory just because we introduced an interaction lagrangian. The reason is that the latter did not include time derivatives and hence did not affect the relation between the canonical momentum π coming with the field ϕ .

Returning to our main line of thought and following our definitions, one straightforwardly finds that, at leading order in λ ,

$$iT_{\text{fi}} = \langle 0 | a_{\vec{p}_1} a_{\vec{p}_2} \left(-i \frac{\lambda}{4!} \right) \int d^4x (\varphi^I(x))^4 a_{\vec{p}_1}^\dagger a_{\vec{p}_2}^\dagger |0\rangle. \quad (5.24)$$

The simplicity of this formula is the reward for our lengthy developments which were aimed at expressing everything (especially U and hence S) in terms of free fields φ^I . These are related to a, a^\dagger in the familiar way, such that one straightforwardly derives

$$iT_{\text{fi}} = -i\lambda(2\pi)^4 \delta^4(p_1 + p_2 - p'_1 - p'_2). \quad (5.25)$$

The δ -distribution enforcing momentum conservation always arises in this context. Hence, it is common to separate it by defining the **invariant matrix element** \mathcal{M}_{fi} through

$$iT_{\text{fi}} = i(2\pi)^4 \delta^4(p_1 + p_2 - p'_1 - p'_2) \mathcal{M}_{\text{fi}}. \quad (5.26)$$

This implies

$$S_{\text{fi}} = \delta_{\text{fi}} + i(2\pi)^4 \delta^4(p_1 + p_2 - p'_1 - p'_2) \mathcal{M}_{\text{fi}}. \quad (5.27)$$

In this language, we have hence found that the invariant matrix element for 2-to-2-scattering in $\lambda\varphi^4$ -theory reads

$$i\mathcal{M}_{\text{fi}} = -i\lambda. \quad (5.28)$$

One represents this graphically by 4 lines meeting at the interaction point. This is then our first Feynman rule:

$$\begin{array}{c} \diagup \\ \diagdown \\ \diagdown \\ \diagup \end{array} = -i\lambda. \quad (5.29)$$

For 2-to-2-scattering in $\lambda\varphi^3$ -theory, one obtains a slightly more complicated result,

$$\begin{array}{c} \diagup \quad \diagdown \\ \text{---} \\ \diagdown \quad \diagup \end{array} \propto \lambda^2. \quad (5.30)$$

We will discuss this later and supplement it by an appropriate factor for the line between the two 3-vertices.

We note that, because of the factor $\delta^4(p_{in} - p_{out})$, our S -matrix element (corresponding to a quantum-mechanical amplitude) is either zero or infinite, depending on whether momentum conservation holds. Hence, in the next section we will have to do some more work, deriving from the singular matrix elements a finite quantity: the scattering cross section.

5.2 Scattering cross section

Consider a so-called **fixed target experiment** (cf. Fig. 5): A beam, which comprises N_B particles of type B , spread over a transverse area F , is directed at a target. The latter is a single particle of type A .

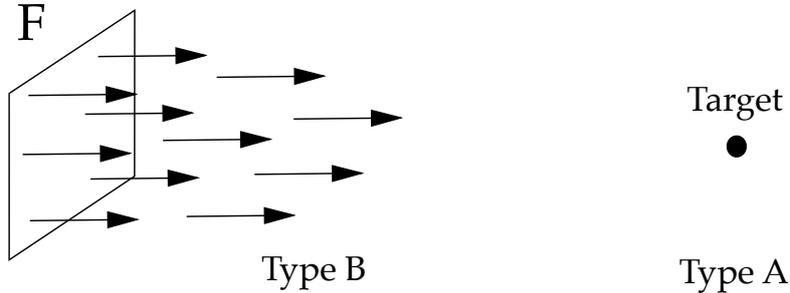


Figure 5: Setup for our idealised fixed-target experiment.

The scattering cross section σ for this experiment is defined by the relations

$$\frac{N_{\text{events}}}{N_B} = \frac{\sigma}{F} \quad \text{or, equivalently,} \quad \sigma = \frac{N_{\text{events}}}{(N_B/F)}. \quad (5.31)$$

Here N_B/F is the (integrated) transverse beam density.⁶

We require localized states and hence consider wave packets:

$$|f_{\vec{p}}\rangle := \int d\tilde{k} f_{\vec{p}}(\vec{k}) |k\rangle, \quad d\tilde{k} \equiv \frac{d^3k}{(2\pi)^3 2k^0}. \quad (5.32)$$

⁶ As you probably recall from the experimental particle physics lectures, in practice one often works with slightly different quantities: One considers a continuous beam, but asks for quantities per unit of time. Thus, one talks about the luminosity, $N_B/(FT) \equiv \mathcal{L}$. Then the rate $R = N_{\text{events}}/T$ is given by $R = \mathcal{L}\sigma$.

Here $f_{\vec{p}}(\vec{k})$ as a function of \vec{k} is peaked near $\vec{k} = \vec{p}$, e.g. $f_{\vec{p}}(\vec{k}) \sim \exp\left(-\alpha|\vec{k} - \vec{p}|^2\right)$.

We choose the following normalization:

$$\langle f_{\vec{p}} | f_{\vec{p}} \rangle = \int d\tilde{k} d\tilde{k}' \overline{f_{\vec{p}}(\tilde{k})} f_{\vec{p}}(\tilde{k}') \langle k | k' \rangle = \int d\tilde{k} |f_{\vec{p}}(\tilde{k})|^2 \equiv 1. \quad (5.33)$$

Next, we check that $|f_{\vec{p}}\rangle$ is indeed localized in \mathbb{R}^3 for an appropriate choice of $f_{\vec{p}}$. Using the by now rather familiar expression for $\varphi^I(x)$ in terms of a and a^\dagger one can readily check that

$$a_{\vec{k}}^\dagger = -i \int d^3x e^{-ikx} \overleftrightarrow{\partial}_0 \varphi^I(x) \Big|_{x^0=0}. \quad (5.34)$$

This gives

$$|f_{\vec{p}}\rangle = \int d\tilde{k} f_{\vec{p}}(\tilde{k}) a_{\tilde{k}}^\dagger |0\rangle = \int d^3x \left\{ \int d\tilde{k} e^{i\vec{k}\vec{x}} (k_0 \varphi^I(\vec{x}) - i \dot{\varphi}^I(\vec{x})) f_{\vec{p}}(\tilde{k}) \right\} |0\rangle. \quad (5.35)$$

The r.h. side is the sum of two terms which are proportional to the following two expressions:

$$\begin{aligned} & \int d^3x \varphi^I(\vec{x}) \left(\int d\tilde{k} e^{i\vec{k}\vec{x}} f_{\vec{p}}(\tilde{k}) k^0 \right) |0\rangle \\ & \int d^3x \dot{\varphi}^I(\vec{x}) \left(\int d\tilde{k} e^{i\vec{k}\vec{x}} f_{\vec{p}}(\tilde{k}) \right) |0\rangle. \end{aligned} \quad (5.36)$$

Now recall that $f_{\vec{p}}(k)$ (and hence also $k^0 f_{\vec{p}}(k)$) is a smooth function, localized in \vec{k} near the origin. Hence the Fourier transforms of both functions will be localized in \vec{x} near $\vec{x} = 0$. Thus, we see that $|f_{\vec{p}}\rangle$ is created by the operators $\varphi^I(x^0 = 0, \vec{x})$ and $\dot{\varphi}^I(x^0 = 0, \vec{x})$ acting on the vacuum in a localized region near $\vec{x} = 0$.

Thus, the state

$$|i\rangle = \int d\tilde{k}_A d\tilde{k}_B f_{\vec{p}_A}(\tilde{k}_A) f_{\vec{p}_B}(\tilde{k}_B) |k_A k_B\rangle \quad (5.37)$$

corresponds, at $t = 0$, to two wave packets localised near $\vec{x} = 0$. The (approximate) momenta are p_A, p_B and the localisation is characterised by the functions f . Figure 6 depicts the situation slightly before $t = 0$, assuming that \vec{p}_B is non-zero while \vec{p}_A vanishes.

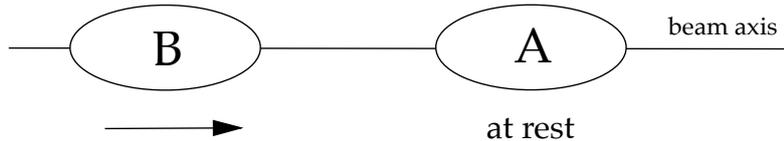


Figure 6: Two wave packets with different momenta before the moment of collision.

Next, we need to account for the transverse spread of incoming particles. In other words, we must allow for a non-zero **impact parameter** \vec{b} , cf. Fig. 7. As in

quantum mechanics, the operator \hat{P} generates shifts. Thus, an incoming particle with $\vec{b} \neq 0$ is given by

$$e^{-i\hat{P}\vec{b}} \int d\vec{k}_B f_{\vec{p}_B}(\vec{k}_B) |k_B\rangle. \quad (5.38)$$

Since $\hat{P}|k_B\rangle = \vec{k}_B|k_B\rangle$, our initial state for given \vec{b} is

$$|i_b\rangle = \int d\vec{k}_A d\vec{k}_B f_{\vec{p}_A}(\vec{k}_A) f_{\vec{p}_B}(\vec{k}_B) e^{-i\vec{k}_B\vec{b}} |k_A k_B\rangle. \quad (5.39)$$

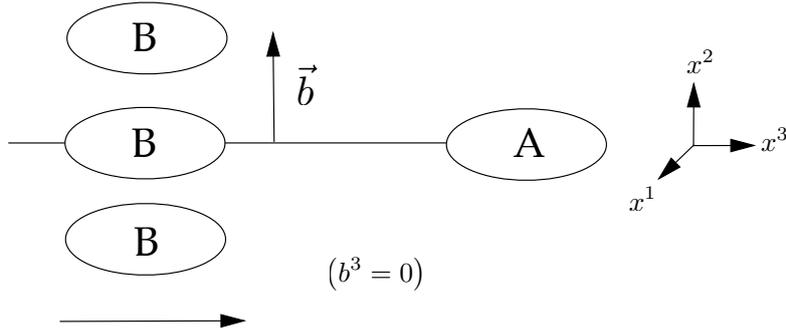


Figure 7: Same as Fig. 6, but now allowing for a non-zero impact parameter.

Starting with $|i_b\rangle$ at $t = -\infty$, when A and B are certainly far apart, we evolve in time through the time region where the particles can make contact and to $t = +\infty$. Then we project on the the desired final state $|p_1 p_2\rangle$. As we learned before, this corresponding amplitude is given characterised by

$$|\langle p_1 p_2 | S | i_b \rangle|^2. \quad (5.40)$$

Summing over a set of particles with different \vec{b} gives

$$N_{\text{events}} = \sum_{\vec{b}} |\langle p_1 p_2 | S | i_b \rangle|^2. \quad (5.41)$$

For a homogeneous, transverse distribution of N_B particles in an area F we can approximate this sum by an integral:

$$N_{\text{events}} = \frac{N_B}{F} \int_F d^2b |\langle p_1 p_2 | S | i_b \rangle|^2. \quad (5.42)$$

According to our definition of the cross section, one furthermore has

$$\sigma(\vec{p}_1, \vec{p}_2) = \frac{N_{\text{events}}}{(N_B/F)} = \int_F d^2b |\langle p_1 p_2 | S | i_b \rangle|^2. \quad (5.43)$$

Clearly, this is too naive as we cannot ask for a specific point of final-state momenta $(\vec{p}_1, \vec{p}_2) \in \mathbb{R}^6$. This also clashes with the finite precision of any realistic detector. Instead, let us define σ for a finite region of phase-space $V_f \subset \mathbb{R}^6$, i.e. write

$$\sigma(V_f) = \int_{V_f} d\vec{p}_1 d\vec{p}_2 \sigma(\vec{p}_1, \vec{p}_2). \quad (5.44)$$

The correctness of the proposed measure $d\tilde{p}_1 d\tilde{p}_2$ follows immediately by considering the free theory ($S = \mathbb{1}$) and integration over the whole phase space: Assuming for simplicity that the two particles are distinguishable, one expects

$$1 = \int d\tilde{p}_1 d\tilde{p}_2 |\langle p_1 p_2 | f_{\vec{p}_A} f_{\vec{p}_B} \rangle|^2 = \int d\tilde{p}_1 d\tilde{p}_2 |f_{\vec{p}_A}(\vec{p}_1)|^2 |f_{\vec{p}_B}(\vec{p}_2)|^2. \quad (5.45)$$

This does indeed hold because of our previously chosen normalisation of f .

All of the above generalises to n final-state particles. Also, we actually want the so-called **differential** cross sections, i.e. (5.44) before integration. This reads

$$d\sigma = \prod_{j=1}^n d\tilde{p}_j \int_F d^2b |\langle p_1 \dots p_n | S | i_b \rangle|^2. \quad (5.46)$$

Now we employ the definition $S_{fi} = \delta_{fi} + i(2\pi)^4 \delta^4(p_f - p_i) \mathcal{M}_{fi}$ and the expression for $|i_b\rangle$ given above. Noting that only the \mathcal{M}_{fi} term contributes unless $f=i$, we find

$$d\sigma = \prod_{j=1}^n d\tilde{p}_j \int d^2b \int d\vec{k}_A d\vec{k}_B f_{\vec{p}_A}(\vec{k}_A) f_{\vec{p}_B}(\vec{k}_B) \int d\vec{k}'_A d\vec{k}'_B \bar{f}_{\vec{p}_A}(\vec{k}'_A) \bar{f}_{\vec{p}_B}(\vec{k}'_B) e^{i\vec{b}(\vec{k}'_B - \vec{k}_B)} |\mathcal{M}_{fi}|^2 (2\pi)^8 \delta^4(p_f - k_i) \delta^4(p_f - k'_i), \quad (5.47)$$

where $p_f = \sum_{j=1}^n p_j$, $k_i = k_A + k_B$ and $k'_i = k'_A + k'_B$. What now follows is a straightforward but slightly painful evaluation of the lengthy expression above. First, we note that

$$\int d^2b e^{i\vec{b}(\vec{k}'_B - \vec{k}_B)} = (2\pi)^2 \delta^2(k'_{B\perp} - k_{B\perp}). \quad (5.48)$$

This allows us to evaluate the transverse parts of the $k'_{A/B}$ integrations as follows:

$$\begin{aligned} & \int d^3k'_A d^3k'_B \delta^4(p_f - k'_i) \delta^2(k'_{B\perp} - k_{B\perp}) \dots \\ &= \int d(k_A^{3'}) d(k_B^{3'}) \delta(p_f^0 - k_i^{0'}) \delta(p_f^3 - k_i^{3'}) \dots \end{aligned} \quad (5.49)$$

After these integrations, one has obviously enforced the equality $k'_{B\perp} = k_{B\perp}$. Less obviously, one also has implemented $k'_{A\perp} = k_{A\perp}$. This follows from $\delta^2(p_{f\perp} - k'_{A\perp} - k'_{B\perp})$ together with $\delta^2(p_{f\perp} - k_{A\perp} - k_{B\perp})$, the latter being hidden in the ellipsis.

Next, one performs the $k_B^{3'}$ integration, giving

$$\int d(k_A^{3'}) \delta(p_f^0 - k_A^{0'} - k_B^{0'}) \dots \quad (5.50)$$

such that now also the relation $k_B^{3'} = p_f^3 - k_A^{3'}$ holds. The result may be written as

$$\int d(k_A^{3'}) \delta\left(p_f^0 - \sqrt{m_A^2 + (\vec{k}'_A)^2} - \sqrt{m_B^2 + (\vec{k}'_B)^2}\right) \dots, \quad (5.51)$$

where $(\vec{k}'_A)^2 = (k_{A\perp})^2 + (k_A^{3'})^2$ and $(\vec{k}'_B)^2 = (k_{B\perp})^2 + (p_f^3 - k_A^{3'})^2$. Performing this last integration gives

$$\left| \frac{k_A^{3'}}{\sqrt{m_A^2 + (\vec{k}'_A)^2}} - \frac{k_B^{3'}}{\sqrt{m_B^2 + (\vec{k}'_B)^2}} \right|^{-1} \dots = \left| \frac{k_A^{3'}}{k_A^{0'}} - \frac{k_B^{3'}}{k_B^{0'}} \right|^{-1} \dots = \frac{1}{|v_A - v_B|} \dots \quad (5.52)$$

Let us furthermore note that, initially, we talked about a fixed-target experiment. This corresponds to $v_A = 0$. However, nothing in our analysis depended on this assumption, so we may as well keep v_A general.

Now, we have completely carried out the $k'_{A,B}$ integrations, implementing the four relations

$$k'_{A\perp} = k_{A\perp}, \quad k'_{B\perp} = k_{B\perp}, \quad k_B^{3'} + k_A^{3'} = p_f^3, \quad k_B^{0'} + k_A^{0'} = p_f^0. \quad (5.53)$$

The last one may also be written as

$$\sqrt{m_B^2 + (k_{B\perp})^2 + (k_B^{3'})^2} + \sqrt{m_A^2 + (k_{A\perp})^2 + (k_A^{3'})^2} = p_f^0. \quad (5.54)$$

Thus, of the four relations in (5.53), we can view the last two as fixing the two variables $k_A^{3'}, k_B^{3'}$. Moreover, two analogous relations hold for k_A^3, k_B^3 . This follows from the so far unused delta-function $\delta^4(p_f - k_A - k_B)$. We have hence enforced that $\vec{k}_A = \vec{k}'_A$ and $\vec{k}_B = \vec{k}'_B$.

Altogether, we now have

$$d\sigma = \prod_j d\tilde{p}_j \int d\tilde{k}_A d\tilde{k}_B \left| f_{\vec{p}_A}(\vec{k}_A) \right|^2 \left| f_{\vec{p}_B}(\vec{k}_B) \right|^2 |\mathcal{M}_{\text{fi}}|^2 \frac{(2\pi)^4 \delta^4(p_f - k_A - k_B)}{4k_A^0 k_B^0 |v_A - v_B|}. \quad (5.55)$$

At this point we may view the two $|f|^2$ as effective δ -functions ensuring $\vec{k}_A = \vec{p}_A, \vec{k}_B = \vec{p}_B$. This gives

$$d\sigma = \frac{|\mathcal{M}_{\text{fi}}|^2}{4p_A^0 p_B^0 |v_A - v_B|} (2\pi)^4 \delta^4(p_f - p_i) \prod_j^n \frac{d^3 p_j}{(2\pi)^3 2p_j^0}. \quad (5.56)$$

The first factor may be rewritten as

$$\frac{|\mathcal{M}_{\text{fi}}|^2}{4p_A^0 p_B^0 |v_A - v_B|} = \frac{|\mathcal{M}_{\text{fi}}|^2}{2w(s, m_A^2, m_B^2)}, \quad (5.57)$$

where

$$w(x, y, z) = \sqrt{x^2 + y^2 + z^2 - 2xy - 2xz - 2yz} \quad \text{and} \quad s = (p_A + p_B)^2. \quad (5.58)$$

We leave the demonstration to the problems. Clearly, this rewriting implies that the factor is invariant under boosts along the x^3 direction. We note that \sqrt{s} is known as the **center-of-mass energy**. The remainder of (5.56) is the so-called **n -particle phase space**.

The highly relativistic case will be the most relevant for us. Thus, let:

$$p_A^0 = m_A; \quad \vec{p}_A = 0; \quad m_A, m_B \ll \sqrt{s}. \quad (5.59)$$

This implies

$$|v_A - v_B| = |v_B| = c = 1 \quad (5.60)$$

and

$$4p_A^0 p_B^0 = 2(p_A + p_B)^2 = 2s. \quad (5.61)$$

In the first equality it has been used that $p_A^2 \approx p_B^2 \approx 0$, which follows directly from (5.59) and the fact that p_B is approximately light-like. Thus, in the highly relativistic case we have:

$$d\sigma = \frac{1}{2s} |\mathcal{M}_{\text{fi}}|^2 dX^{(n)}; \quad dX^{(n)} = (2\pi)^4 \delta^4(p_f - p_i) d\tilde{p}_1 \dots d\tilde{p}_n. \quad (5.62)$$

5.3 2-particle phase-space and a simple example

Consider specifically 2-to-2 scattering in $\lambda\varphi^4$ -theory and focus on the phase space:

$$\int dX^{(2)} = \int (2\pi)^4 \delta^4(p_1 + p_2 - p_A - p_B) \frac{d^3 p_1}{(2\pi)^3 2p_1^0} \frac{d^3 p_2}{(2\pi)^3 2p_2^0}. \quad (5.63)$$

The $d^3 p_1$ -integration may be trivially performed, giving

$$\int \frac{d^3 p_2}{(2\pi)^2 4p_1^0 p_2^0} \delta(p_1^0 + p_2^0 - \sqrt{s}) = \int \frac{d^3 p_2}{(2\pi)^2 4|\vec{p}_2|^2} \delta(2|\vec{p}_2| - \sqrt{s}). \quad (5.64)$$

In the second step, we used $\vec{p}_1 = -\vec{p}_2$, which corresponds to choosing the center-of-mass frame. In addition, we assumed $p_2^2 = (p_2^0)^2 - \vec{p}_2^2 = m^2 \simeq 0$, justified by the highly relativistic limit.

Next, we switch to spherical coordinates:

$$d^3 p_2 = d\Omega |\vec{p}_2|^2 d|\vec{p}_2|; \quad d\Omega = d\varphi \sin\theta d\theta, \quad (5.65)$$

and perform the $d|\vec{p}_2|$ integration only. In other words, we keep our result differential concerning the angular distribution:

$$\int d|\vec{p}_2| \delta(2|\vec{p}_2| - \sqrt{s}) \frac{d\Omega}{16\pi^2} = \frac{d\Omega}{32\pi^2}. \quad (5.66)$$

Combining this with the invariant matrix element squared and the prefactor $1/2s$, we obtain the **differential cross section**

$$\frac{d\sigma}{d\Omega} = \frac{|\mathcal{M}_{\text{fi}}|^2}{64\pi^2 s} = \frac{\lambda^2}{64\pi^2 s}. \quad (5.67)$$

We see that no direction is preferred, which is expected since our particles carry no spin and also our interaction term involves no directional information (e.g. derivatives acting on the fields etc.). Moreover, the factor λ^2/s could have been argued without any calculations: The scattering amplitude is proportional to λ and therefore λ has to appear quadratically in the differential cross section. Cross sections have units of $[\text{Length}]^2$ corresponding to $[\text{Energy}]^{-2}$. Hence, we have to divide λ^2 by the energy scale squared: s . A possible mass dependence is irrelevant as $m \ll \sqrt{s}$ by assumption. The numerical coefficient, however, required the detailed analysis just performed.

6 LSZ-Formalism

This goes back to the 1955 paper by Lehmann, Symanzik, and Zimmermann. Our presentation is similar to that by Peskin and Schröder [1] which in turn is similar to Weinberg's approach [4]. But it differs from many other books and the reader is invited to compare the discussions of this non-trivial subject by different authors.

The general idea is to establish a relation between S-matrix-elements and correlation functions (also known as Green's functions). Concretely:

$$\underbrace{\langle p'_1 \dots p'_n | p_1 \dots p_m \rangle_{\text{in}}}_{\text{needed for cross section}} \longleftrightarrow \underbrace{\langle 0 | T \varphi(x_1) \dots \varphi(x_{n+m}) | 0 \rangle}_{\text{easily calculable in pert. theory}} .$$

This will improve our naive discussion of the last chapter not only by providing a systematic treatment extendible to higher orders in perturbation theory. Even more importantly, it will correct our simplifying assumption that incoming and outgoing particles can be identified with Fock space basis states. The relation will turn out to be more complicated.

6.1 Spectral density and Z-factors

In the following we will use the Heisenberg picture. So it will be convenient to write the corresponding objects, the time-dependent field operators and the time-independent states, without any index:

$$\varphi(x) = e^{iHt} \varphi_S(\vec{x}) e^{-iHt} ; \quad H = H_0 + H_{\text{int}} ; \quad |\psi\rangle = |\psi_S(t=0)\rangle . \quad (6.1)$$

The index S stands for Schrödinger picture.

Consider the correlation function:

$$\langle 0 | \varphi(x) \varphi(y) | 0 \rangle . \quad (6.2)$$

This may be interpreted as the amplitude for a particle to propagate from \vec{y} at time y^0 to \vec{x} at time x^0 . For a free field φ_0 with mass m_0 we have:

$$\begin{aligned} \langle 0 | \varphi_0(x) \varphi_0(y) | 0 \rangle &= \int d\vec{p} d\vec{q} \langle 0 | \hat{a}_{\vec{p}} \hat{a}_{\vec{q}}^\dagger | 0 \rangle e^{-ipx+iqy} \\ &= \int \frac{d^3p}{(2\pi)^3 2p^0} e^{-ip(x-y)} = \int \frac{d^4p}{(2\pi)^3} e^{-ip(x-y)} \delta(p^2 - m_0^2) \Theta(p^0) \\ &\equiv D(x-y, m_0^2) , \end{aligned} \quad (6.3)$$

where Θ denotes the Heaviside step function.

Now we consider the general case, where we can write

$$\langle 0 | \varphi(x) \varphi(y) | 0 \rangle = \sum_{\alpha} \langle 0 | \varphi(x) | \alpha \rangle \langle \alpha | \varphi(y) | 0 \rangle . \quad (6.4)$$

Here the sum includes *all* states, also multi-particle states. The parameter α is in part continuous. It labels different (discrete) types of intermediate states but also

the continuous parameters associated with each such state, e.g. the total momentum of a state with a fixed number of particles. The ‘sum’ is hence not a proper discrete sum but should rather be understood symbolically.

By using the two relations

$$\varphi(x) = e^{i\hat{p}x}\varphi(0) e^{-i\hat{p}x}, \quad e^{-i\hat{p}x}|\alpha\rangle = e^{-ip_\alpha x}|\alpha\rangle \quad (6.5)$$

we obtain

$$\begin{aligned} \langle 0|\varphi(x)\varphi(y)|0\rangle &= \sum_{\alpha} e^{-ip_{\alpha}(x-y)}|\langle 0|\varphi(0)|\alpha\rangle|^2 \\ &= \int d^4q \sum_{\alpha} e^{-iq(x-y)}|\langle 0|\varphi(0)|\alpha\rangle|^2 \delta^4(q-p_{\alpha}) \\ &= \int \frac{d^4q}{(2\pi)^3} e^{-iq(x-y)}\rho(q). \end{aligned} \quad (6.6)$$

In the last line we used the definition

$$\rho(q) \equiv (2\pi)^3 \sum_{\alpha} \delta^4(q-p_{\alpha})|\langle 0|\varphi(0)|\alpha\rangle|^2. \quad (6.7)$$

Note that $\rho(q)$ is manifestly $\text{SO}^+(1,3)$ -invariant and vanishes for $q^0 < 0$. This allows us to introduce the **spectral density** $\sigma(q^2)$ as follows:

$$\rho(q) \equiv \Theta(q^0)\sigma(q^2). \quad (6.8)$$

The spectral density σ quantifies the contribution of the intermediate states $|\alpha\rangle$ with $p_{\alpha}^2 = q^2$ to the correlation function we started from.

Further rewriting yields:

$$\begin{aligned} \langle 0|\varphi(x)\varphi(y)|0\rangle &= \int_0^{\infty} d(M^2) \int \frac{d^4q}{(2\pi)^3} e^{-iq(x-y)}\delta(q^2-M^2)\Theta(q^0)\sigma(M^2) \\ &= \int_0^{\infty} d(M^2) D(x-y, M^2)\sigma(M^2). \end{aligned} \quad (6.9)$$

It can be seen easily that by choosing $\sigma(q^2) = \delta(q^2 - m_0^2)$ we obtain the result (6.3) of the free-field case.

More generally, the spectral density approximately takes the form displayed in Fig. 8. Note that we assume that the vacuum does not contribute as an intermediate state, $\langle 0|\varphi(x)|0\rangle$. This can always be ensured by a redefinition $\varphi \rightarrow \varphi + \text{const}$. Thus, we may write

$$\sigma(q^2) = Z \delta(q^2 - \underbrace{m^2}_{\neq m_0^2}) + \dots \quad (6.10)$$

where the Z is a so far unknown normalization factor. We get further non-zero contributions, indicated by the ellipsis, for $q^2 > M_t^2$. Here, as shown in the figure,

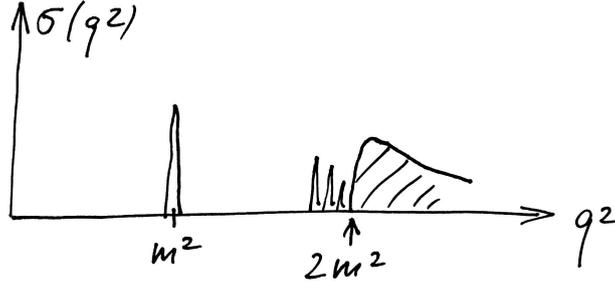


Figure 8: Qualitative behaviour of the spectral density. The peaks just below $(2m)^2$ correspond to multi-particle bound states or resonances. They start at the so-called **multi-particle-threshold** M_t^2 .

the multi-particle threshold M_t^2 lies in general below $2m^2$. The reason is that 2-particle bound states have a mass that is smaller than then the minimal mass $2m^2$ of two particles separated by a large distance.

Given this understanding, we now split the formal sum \sum_{α} into a single-particle part (involving the integration over the momentum) and a separate multi-particle contribution (denoted by an ellipsis):

$$\sum_{\alpha} |\alpha\rangle \langle \alpha| = \int d\tilde{p} |p\rangle \langle p| + \dots \quad (6.11)$$

We then obtain

$$\begin{aligned} \langle 0 | \varphi(x) \varphi(y) | 0 \rangle &= \int d\tilde{p} \langle 0 | \varphi(x) | p \rangle \langle p | \varphi(y) | 0 \rangle + \dots \\ &= \int d\tilde{p} e^{-ip(x-y)} \underbrace{|\langle 0 | \varphi(0) | p \rangle|^2}_{\equiv Z} + \dots \\ &= D(x-y, m^2) Z + \dots \end{aligned} \quad (6.12)$$

In the last step we used the fact that Z does not depend on p . This is clear since one may write $|p\rangle = \hat{\Lambda} |p'\rangle$ and $\langle 0| = \langle 0| \hat{\Lambda}^\dagger$, noting also that $\hat{\Lambda}^\dagger \varphi(0) \hat{\Lambda} = \varphi(0)$. Here we assumed that Lorentz covariance, as introduced in the free-field case, continues to hold in the interacting theory.

Altogether, we have shown that

$$\langle 0 | \varphi(x) \varphi(y) | 0 \rangle = Z D(x-y, m^2) + \underbrace{\int_{M_t^2}^{\infty} d(M^2) \sigma(M^2) D(x-y, M^2)}_{\text{multi-particle contribution}}. \quad (6.13)$$

Subtracting the same equation with $x \leftrightarrow y$ yields

$$\langle 0 | [\varphi(x), \varphi(y)] | 0 \rangle = Z \Delta(x-y, m^2) + \int_{M_t^2}^{\infty} d(M^2) \sigma(M^2) \Delta(x-y, M^2), \quad (6.14)$$

with

$$\Delta(x - y, M^2) \equiv \langle 0 | [\varphi_0(x), \varphi_0(y)] | 0 \rangle . \quad (6.15)$$

Here the free theory, obviously, is taken to have mass M .

Now we apply $\left. \frac{\partial}{\partial y^0} \right|_{y^0=x^0}$ to (6.15). Since $\dot{\varphi} = \pi$, we find

$$[\varphi_0(x^0, \vec{x}), \pi_0(x^0, \vec{y})] = i \delta^3(\vec{x} - \vec{y}) \quad (6.16)$$

on the r.h. side. This implies

$$\left. \frac{\partial}{\partial y^0} \Delta(x - y, M^2) \right|_{y^0=x^0} = i \delta^3(\vec{x} - \vec{y}) . \quad (6.17)$$

Next, we apply $\left. \frac{\partial}{\partial y^0} \right|_{y^0=x^0}$ to (6.14) and observe that (6.16) also holds for interacting fields by the very definition of our quantization procedure:

$$[\varphi(x^0, \vec{x}), \pi(x^0, \vec{y})] = i \delta^3(\vec{x} - \vec{y}) . \quad (6.18)$$

Thus, we obtain

$$1 = Z + \int_{M_t^2}^{\infty} d(M^2) \sigma(M^2) . \quad (6.19)$$

This implies that $Z \leq 1$ and $Z = 1$ precisely for the free theory. The size of $1 - Z$ accounts for the overlap of $\varphi(0) | 0 \rangle$ with multi-particle states.

Finally, we may use (6.13) to build the expression $T\varphi(x)\varphi(y)$ rather than $[\varphi(x), \varphi(y)]$ of (6.14). The result is

$$\langle 0 | T\varphi(x)\varphi(y) | 0 \rangle = Z D_F(x - y, m^2) + \int_{M_t^2}^{\infty} d(M^2) \sigma(M^2) D_F(x - y, M^2) \quad (6.20)$$

The Feynman propagator

$$D_F(x - y, m_0^2) \equiv \langle 0 | T\varphi_0(x)\varphi_0(y) | 0 \rangle , \quad (6.21)$$

which appears on the r.h. side, will be of particular interest to us in the following sections.

6.2 LSZ reduction formula

Now we will relate time-ordered correlation functions to scattering amplitudes. We start by considering the Fourier transform

$$\begin{aligned} & \int d^4x e^{ipx} \langle 0 | T\varphi(x)\varphi(z_1) \dots \varphi(z_n) | 0 \rangle \\ &= \left(\int_{-\infty}^{T_-} dx^0 + \int_{T_-}^{T_+} dx^0 + \int_{T_+}^{\infty} dx^0 \right) \left(\int d^3x e^{ipx} \langle 0 | T\varphi(x)\varphi(z_1) \dots \varphi(z_n) | 0 \rangle \right) \\ &\equiv \text{(I)} + \text{(II)} + \text{(III)} , \end{aligned} \quad (6.22)$$

which we split into three terms, as defined above. We will view this expression as a function of the complex variable p^0 . Our interest will be in its **pole structure** in p^0 . In particular, we claim that there is a pole at $p^0 = \sqrt{\vec{p}^2 + m^2} = \omega_{\vec{p}}$ and we determine its residue. First, we focus on integration region III and hence on the third term:

$$\begin{aligned} \int_{T_+}^{\infty} dx^0 (\dots) &= \int_{T_+}^{\infty} dx^0 \int d^3x e^{ipx} \sum_{\alpha} \langle 0 | \varphi(x) | \alpha \rangle \langle \alpha | T\varphi(z_1) \dots \varphi(z_n) | 0 \rangle \quad (6.23) \\ &= \sum_{\alpha} \int_{T_+}^{\infty} dx^0 \int d^3x e^{ix^0(p^0 - q_{\alpha}^0) - i\vec{x}(\vec{p} - \vec{q}_{\alpha})} \langle 0 | \varphi(0) | \alpha \rangle \langle \alpha | T\varphi(z_1) \dots \varphi(z_n) | 0 \rangle . \end{aligned}$$

A pole in p^0 will arise if the oscillating exponent vanishes, i.e. if $p^0 - q_{\alpha}^0 = 0$. Thus, a pole at $p^0 = \omega_{\vec{p}}$ can only come from 1-particle-states. Indeed, the d^3x integration enforces $\vec{q}_{\alpha} = \vec{p}$, such that the 4-vector q_{α} must obey $q_{\alpha}^2 = m^2$. As a result, we may replace $\sum_{\alpha} | \alpha \rangle \langle \alpha | \rightarrow \int d\tilde{q}_{\alpha} | q_{\alpha} \rangle \langle q_{\alpha} |$.

We now defined the symbol \sim to mean ‘equal up to finite terms’ or, in other words, ‘the coefficients of the poles agree’. With this, the previous equation may be continued as

$$\begin{aligned} \dots &\sim \int_{T_+}^{\infty} dx^0 \int d^3x \int d\tilde{q}_{\alpha} e^{ix^0(p^0 - q_{\alpha}^0) - i\vec{x}(\vec{p} - \vec{q}_{\alpha})} \langle 0 | \varphi(0) | q \rangle \langle q | T\varphi(z_1) \dots \varphi(z_n) | 0 \rangle \\ &\sim \frac{\sqrt{Z}}{2\omega_{\vec{p}}} \int_{T_+}^{\infty} dx^0 e^{ix^0(p^0 - \omega_{\vec{p}})} \langle p | T\varphi(z_1) \dots \varphi(z_n) | 0 \rangle \end{aligned} \quad (6.24)$$

We evaluate this integral assuming that p^0 has a small, positive imaginary part. This is legitimate since we are treating the whole expression as an analytic function of p^0 and we only care about the residue of the pole at $p^0 = \omega_{\vec{p}}$. The result is

$$\sim \frac{\sqrt{Z}}{2\omega_{\vec{p}}} \cdot \frac{1}{i(p^0 - \omega_{\vec{p}})} \left(0 - e^{ix^0(p^0 - \omega_{\vec{p}})} \Big|_{x^0=T_+} \right) \langle p | T\varphi(z_1) \dots \varphi(z_n) | 0 \rangle . \quad (6.25)$$

Now the pole is manifest and the remaining exponent can be dropped since it equals unity at the pole. Moreover, we have $p^2 - m^2 = (p^0 + \omega_{\vec{p}})(p^0 - \omega_{\vec{p}}) \simeq 2\omega_{\vec{p}}(p^0 - \omega_{\vec{p}})$, such that we eventually find

$$\sim \frac{i\sqrt{Z}}{p^2 - m^2} \langle p | T\varphi(z_1) \dots \varphi(z_n) | 0 \rangle . \quad (6.26)$$

This is our final, maximally simplified expression displaying the desired pole-structure at $p^0 = \omega_{\vec{p}}$.

Now we focus on integration region I. It is easy to write the analogue of (6.23):

$$\begin{aligned} \int_{-\infty}^{T_-} dx^0 (\dots) &= \int_{-\infty}^{T_-} dx^0 \int d^3x e^{ipx} \sum_{\alpha} \langle 0|T\varphi(z_1)\dots\varphi(z_n)|\alpha\rangle\langle\alpha|\varphi(x)|0\rangle \quad (6.27) \\ &= \sum_{\alpha} \int_{-\infty}^{T_-} dx^0 \int d^3x e^{ix^0(p^0+q_{\alpha}^0)-i\vec{x}(\vec{p}+q_{\alpha}^{\vec{}})} \langle 0|T\varphi(z_1)\dots\varphi(z_n)|\alpha\rangle\langle\alpha|\varphi(0)|0\rangle. \end{aligned}$$

We see that the change $\langle 0|\varphi(x)|\alpha\rangle \rightarrow \langle\alpha|\varphi(x)|0\rangle$ has led to a sign flip of q_{α} in the exponent. Now the coefficient of x^0 in the exponent is always positive. The exponential oscillates at $p^0 \rightarrow \omega_{\vec{p}}$, so no pole arises.

Finally, integration region II is finite, hence its contribution is analytic in p^0 , so no pole arises here either.

We have arrived at an important **preliminary result**: We managed to trade $\varphi(x)$ in (6.22) for an outgoing particle $\langle p|$ in (6.26).

A completely analogous calculation can be performed for a negative sign of the exponent in the Fourier transformation:

$$\int d^4x e^{-ipx} \langle 0|T\varphi(x)\varphi(z_1)\dots\varphi(z_n)|0\rangle \sim \frac{i\sqrt{Z}}{p^2 - m^2} \langle 0|T\varphi(z_1)\dots\varphi(z_n)|p\rangle. \quad (6.28)$$

It allows us to trade $\varphi(x)$ for an incoming particle.

Finally we need to be able to do several such manipulations at once, e.g.

$$\begin{aligned} \int d^4x_1 \int d^4x_2 e^{ip_1x_1+ip_2x_2} \langle 0|T\varphi(x_1)\varphi(x_2)\varphi(z_1)\dots\varphi(z_n)|0\rangle \\ \sim \frac{i\sqrt{Z}}{p_1^2 - m_1^2} \cdot \frac{i\sqrt{Z}}{p_2^2 - m_2^2} \langle p_1p_2|T\varphi(z_1)\dots\varphi(z_n)|0\rangle. \end{aligned} \quad (6.29)$$

There are now many fields and many integration regions. For illustration, let us display only two fields and focus on the integration region where both x_i^0 are large. Then our derivation will go through if one may treat the fields separately, roughly as follows:

$$\langle 0|T\varphi(x_1)\varphi(x_2)\dots|0\rangle = \sum_{\alpha,\beta} \langle 0|\varphi(x_1)|\alpha\rangle \langle 0|\varphi(x_2)|\beta\rangle \langle\alpha,\beta|\dots|0\rangle \quad (6.30)$$

This will hold if \vec{x}_1, \vec{x}_2 are always far apart in the region where $x_1^0, x_2^0 \rightarrow \infty$. For a double Fourier-transform in x_1, x_2 this is certainly not true. Hence we need wave packets:

$$\int d^4x e^{ipx} \rightarrow \int d^4x \underbrace{\int d\tilde{k} f_{\vec{p}}(\vec{k}) e^{i\tilde{k}x}}_{\text{function of } \vec{x} \text{ which is localized near zero at } x^0 \text{ and correspondingly in other regions of space at other times.}} \quad (6.31)$$

Thus, the overall structure of the derivation is as follows: Convolute $\langle 0|T\varphi(z_1)\dots\varphi(z_n)|0\rangle$ with appropriate wave packets. Show that the result is some

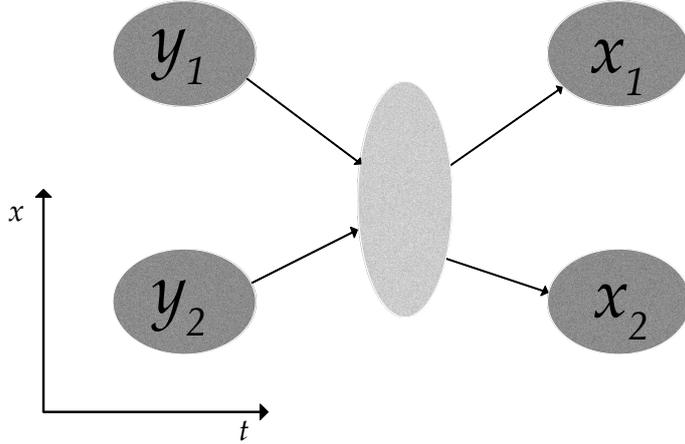


Figure 9: Visualization of wave packets governing the integration with different field operators. For example, at late times $\varphi(x_2)$ only contributes for \vec{x}_2 in the appropriate shaded region.

linear combination of poles. The residues are the desired matrix elements. Take the limit $f_{\vec{p}}(\vec{k}) \rightarrow \delta(\vec{k} - \vec{p})$ at the end. Justifying this last limit rigorously is non-trivial but if it works, one obtains the following **LSZ-Reduction Formula**:

$$\begin{aligned} & \prod_{l=1}^n \int d^4x_l e^{ip_l x_l} \prod_{j=1}^m \int d^4y_j e^{-ik_j y_j} \langle 0 | T \varphi(x_1) \dots \varphi(x_n) \varphi(y_1) \dots \varphi(y_m) | 0 \rangle \\ & \sim \prod_{l=1}^n \frac{i\sqrt{Z}}{p_l^2 - m_l^2} \prod_{j=1}^m \frac{i\sqrt{Z}}{k_j^2 - m_j^2} \underbrace{\langle p_1 \dots p_n | k_1 \dots k_m \rangle_{\text{in}}}_{\equiv S_{\text{fi}}} \end{aligned} \quad (6.32)$$

Crucially, the notation $|k_1 k_2\rangle_{\text{in}}$ characterizes a state in the Heisenberg picture of the fully interaction theory which describes two incoming wave-packets of particles $|k_1\rangle$, $|k_2\rangle$ (separated in \vec{x}) at $x^0 \rightarrow -\infty$. An analogous definition applies to $\langle p_1 \dots p_n |_{\text{out}}$.

We will calculate the S_{fi} by Fourier-transforming time-ordered correlation functions and extracting the relevant residue.

6.3 Calculating time-ordered correlation functions

As explained, we are interested in calculating

$$\langle 0 | T \varphi(x_1) \dots \varphi(x_n) | 0 \rangle \quad (6.33)$$

for any n . Let us focus $n = 2$ and for notational simplicity set $x_1 = x'$, $x_2 = x$. Our argument will go through identically for $n > 2$.

Let us assume $x'^0 > x^0$ without loss of generality and suppress \vec{x}, \vec{x}' , simply to keep notation at a minimum:

$$\langle 0 | T \varphi(x') \varphi(x) | 0 \rangle = \langle 0 | \varphi(x') \varphi(x) | 0 \rangle = \langle 0 | \varphi(t') \varphi(t) | 0 \rangle. \quad (6.34)$$

We may transit from the Heisenberg to the interaction picture as follows:

$$\begin{aligned}
\langle 0 | \varphi(t') \varphi(t) | 0 \rangle &= \langle 0 | e^{iHt'} \varphi(0) e^{-iH(t'-t)} \varphi(0) e^{-iHt} | 0 \rangle \\
&= \langle 0 | \underbrace{e^{iHt'} e^{-iH_0 t'}}_{U(0, t')} \underbrace{e^{iH_0 t'} \varphi(0) e^{-iH_0 t'}}_{\varphi_I(t')} \underbrace{e^{iH_0 t'} e^{-iH(t'-t)} e^{-iH_0 t}}_{U(t', t)} \underbrace{e^{iH_0 t} \varphi(0) e^{-iH_0 t}}_{\varphi_I(t)} \underbrace{e^{iH_0 t} e^{-iHt}}_{U(t, 0)} | 0 \rangle \\
&= \langle 0 | U(0, t') \varphi_I(t') U(t', t) \varphi_I(t) U(t, 0) | 0 \rangle \\
&= \langle 0 | U(0, \infty) U(\infty, t') \varphi_I(t') U(t', t) \varphi_I(t) U(t, -\infty) U(-\infty, 0) | 0 \rangle .
\end{aligned} \tag{6.35}$$

Here $U(t, -\infty)$ is understood as $U(t, T)$, the unitary operator evolving states in the interaction picture, in the limit $T \rightarrow -\infty$.

This last limit only makes sense if interactions are adiabatically switched off at $t \rightarrow \pm\infty$. This can be realised by replacing the Hamiltonian according to

$$H_0 + H_{int} \rightarrow H_0 + f(t) H_{int} , \tag{6.36}$$

with a smooth function $f(t)$ that vanishes at small and large times. Note that quantum-mechanical time evolution is still unitary – quantum mechanics allows for such fundamentally time-dependent Hamiltonians.

By adiabaticity, the interacting vacuum now evolves into the *free vacuum* in the limit $t \rightarrow \pm\infty$. This free vacuum is denoted by $|0\rangle_0$. It is by definition the state annihilated by all $a_{\vec{p}}$. The $a_{\vec{p}}$ are related to φ_I exactly as in the free theory.⁷

With this we can conclude that

$$U(t, -\infty) U(-\infty, 0) |0\rangle = U(t, -\infty) |0\rangle_0 \langle 0 | U(-\infty, 0) |0\rangle , \tag{6.37}$$

and furthermore

$$\langle 0 | T \varphi(t') \varphi(t) | 0 \rangle = \frac{{}_0\langle 0 | U(\infty, t') \varphi_I(t') U(t', t) \varphi_I(t) U(t, -\infty) |0\rangle_0}{{}(\langle 0 | U(0, \infty) |0\rangle_0 \langle 0 | U(-\infty, 0) |0\rangle)^{-1}} . \tag{6.38}$$

Here by unitarity the denominator is a product of two phases. Obviously, for phases the inverse is identical to the complex conjugate, $(\dots)^{-1} = \overline{(\dots)}$. Hence the denominator becomes

$$\begin{aligned}
\overline{{}_0\langle 0 | U(0, \infty) |0\rangle_0 \langle 0 | U(-\infty, 0) |0\rangle} &= {}_0\langle 0 | U(\infty, 0) |0\rangle \langle 0 | U(0, -\infty) |0\rangle_0 \\
&= {}_0\langle 0 | U(\infty, -\infty) |0\rangle_0 .
\end{aligned} \tag{6.39}$$

The next step is to remember the formula for $U(t, t')$ obtained in 5.15 and apply it to both numerator and denominator of the previous expression. For the numerator, we may write the time ordering symbol in front of the whole expression:

$$\begin{aligned}
{}_0\langle 0 | T \exp \left(-i \int_{t'}^{\infty} H_{int}^I d\tau \right) \varphi_I(t') \\
\exp \left(-i \int_t^{t'} H_{int}^I d\tau \right) \varphi_I(t) \exp \left(-i \int_{-\infty}^t H_{int}^I d\tau \right) |0\rangle_0 .
\end{aligned} \tag{6.40}$$

⁷ Recall that the index I stands for ‘interaction picture’, *not* for ‘interacting’.

Since under the time ordering symbol the order of operators does not matter – they will be ordered by T anyway – we can combine the three exponents into one. For the denominator, we simply apply 5.15 – no further manipulations are needed.

With these results and generalizing to $n \geq 2$ we finally get

$$\begin{aligned} & \langle 0|T\varphi(x_1)\cdots\varphi(x_n)|0\rangle \\ &= \frac{{}_0\langle 0|T\varphi_I(x_1)\cdots\varphi_I(x_n)\exp\left(-i\int_{-\infty}^{\infty}d\tau H_{int}(\varphi_I(\tau,\vec{x}))\right)|0\rangle_0}{{}_0\langle 0|T\exp\left(-i\int_{-\infty}^{\infty}d\tau H_{int}(\varphi_I(\tau,\vec{x}))\right)|0\rangle_0}. \end{aligned} \quad (6.41)$$

We may now take the limit $f(t) \rightarrow 1$ and return to our Poincare invariant theory.

Our last formula represents enormous progress: We can now evaluate time-ordered correlation functions of the interacting theory using just free-field commutation relations.

7 Wick-Theorem and Feynman Rules

7.1 Time ordering and normal ordering

As we have just seen, we need to work out expressions like

$${}_I\langle 0|T\varphi_I(x_1)\cdots\varphi_I(x_n)\exp\left(i\int d^4x\mathcal{L}_{int}(\varphi_I(x))\right)|0\rangle_I. \quad (7.1)$$

Here we used $-i\int d\tau H_{int} = i\int d^4x\mathcal{L}_{int}$ to give our expression a manifestly covariant form.

We now want to study these expressions in more detail. As this whole section will mostly treat free fields, we drop the index I for ‘interaction picture’ and simply write

$$\varphi_I \rightarrow \varphi \quad |0\rangle_0 \rightarrow |0\rangle. \quad (7.2)$$

Since \mathcal{L}_{int} is a polynomial in φ , we can expand the exponential and reduce our last expression to a sum of *free-field correlation functions*

$$\langle 0|T\varphi(x_1)\cdots\varphi(x_m)|0\rangle. \quad (7.3)$$

Here $m \geq n$ and, we emphasize this again, φ are now free or, equivalently, interaction-picture fields. We have already assigned a name to the case $m = 2$,

$$\langle 0|T\varphi(x)\varphi(y)|0\rangle = D_F(x-y) = \text{Feynman propagator}, \quad (7.4)$$

and we could easily evaluate it since we know $\langle 0|\varphi(x)\varphi(y)|0\rangle$.

To be able to generalize to the multi-field case, it is useful to split the field in a creation part φ^c and an annihilation part φ^a

$$\varphi(x) = \int d\tilde{k} \left(a_{\vec{k}} e^{-ikx} + a_{\vec{k}}^\dagger e^{ikx} \right) = \varphi^a(x) + \varphi^c(x). \quad (7.5)$$

Next we define the **normal-ordered** form of any operator:

$$: \left(a_{\vec{k}_1}^- a_{\vec{k}_2}^\dagger a_{\vec{k}_3}^\dagger a_{\vec{k}_4}^- \cdots a_{\vec{k}_n}^\dagger \right) :, \equiv a_{\vec{k}_2}^\dagger a_{\vec{k}_3}^\dagger \cdots a_{\vec{k}_n}^\dagger a_{\vec{k}_1}^- a_{\vec{k}_4}^- \cdots . \quad (7.6)$$

The double-dots on both sides of the expression say that it is normal-ordered. The latter is defined simply by placing all creation operators to the left of all annihilation operators. Since φ is a linear combination of a, a^\dagger this definition extends to any product of φ . In particular one gets

$$: \varphi^a(x) \varphi^c(y) : = \varphi^c(y) \varphi^a(x) \quad (7.7)$$

Note the following two facts: First, for any operator \hat{O} that is a polynomial or series in a, a^\dagger without a constant term, we find

$$\langle 0 | : \hat{O} : | 0 \rangle = 0. \quad (7.8)$$

Second, our prescription for dropping the vacuum energy amounts to saying:

$$H_0 = : \frac{1}{2} \int d^3x \left(\pi^2 + (\vec{\nabla}\varphi)^2 + m^2\varphi^2 \right) : \quad (7.9)$$

The product of two fields, $\varphi(x)\varphi(y)$, and the corresponding normal-ordered expression, $: \varphi(x)\varphi(y) :$, differ only by a number. (To be precise, by this we mean it is a number multiplied by the identity operator.) This is clear since the commutator of two creation/annihilation operators is just a number:

$$\begin{aligned} \varphi(x)\varphi(y) &= (\varphi_x^a + \varphi_x^c)(\varphi_y^a + \varphi_y^c) = \varphi_x^a\varphi_y^a + \varphi_y^c\varphi_x^a + \varphi_x^c\varphi_y^a + \varphi_x^c\varphi_y^c + \underbrace{[\varphi_x^a, \varphi_y^c]}_{\text{"number"}} \\ &= : \varphi(x)\varphi(y) : + \text{"number"}. \end{aligned} \quad (7.10)$$

Sandwiching this in the vacuum state, we find

$$\langle 0 | \varphi(x)\varphi(y) | 0 \rangle = \text{"number"} \quad (7.11)$$

and hence

$$\varphi(x)\varphi(y) = : \varphi(x)\varphi(y) : + \langle 0 | \varphi(x)\varphi(y) | 0 \rangle. \quad (7.12)$$

We can repeat this whole derivation starting from the time-ordered product of two fields. The result is

$$T\varphi(x)\varphi(y) = : \varphi(x)\varphi(y) : + \langle 0 | T\varphi(x)\varphi(y) | 0 \rangle \quad (7.13)$$

Finally, we introduce the convenient notation

$$\langle 0 | T\varphi(x)\varphi(y) | 0 \rangle = \overline{\varphi(x)\varphi(y)}, \quad (7.14)$$

calling this a **contraction**. We may then write

$$T\varphi(x)\varphi(y) = : \varphi(x)\varphi(y) : + \overline{\varphi(x)\varphi(y)} \quad (7.15)$$

7.2 Wick theorem

The statement of the **Wick theorem** is very simple:

$$T\varphi(x_1)\cdots\varphi(x_n) = : \varphi(x_1)\cdots\varphi(x_n) : + \text{all contractions of } : \varphi(x_1)\cdots\varphi(x_n) : . \quad (7.16)$$

Here, on the r.h. side, we have to sum over all terms that arise by contracting one or more pairs of fields in the original expression. Contracting two fields means removing them from the product and instead multiplying by the vacuum expectation value of the corresponding time-ordered product.

To work out an example, it is convenient to introduce the shorthand notation $\varphi_i \equiv \varphi(x_i)$. With this, we have

$$\begin{aligned} T\varphi_1\varphi_2\varphi_3\varphi_4 &= : \varphi_1\varphi_2\varphi_3\varphi_4 : \\ &+ (: \overline{\varphi_1\varphi_2}\varphi_3\varphi_4 : + 5 \text{ analogous terms}) \\ &+ (: \overline{\varphi_1\varphi_2}\overline{\varphi_3\varphi_4} : + : \overline{\varphi_1\varphi_2\varphi_3}\varphi_4 : + : \overline{\varphi_1\varphi_2\varphi_3\varphi_4} :). \end{aligned} \quad (7.17)$$

The normal ordering in the last term can be dropped since, due to the contractions, we are just dealing with numbers. Moreover, we can identify the doubly contracted terms with products of two Feynman propagators:

$$\begin{aligned} &(: \overline{\varphi_1\varphi_2}\overline{\varphi_3\varphi_4} : + : \overline{\varphi_1\varphi_2\varphi_3}\varphi_4 : + : \overline{\varphi_1\varphi_2\varphi_3\varphi_4} :) \\ &= D_F(x_1 - x_2)D_F(x_3 - x_4) + D_F(x_1 - x_4)D_F(x_2 - x_3) \\ &+ D_F(x_1 - x_3)D_F(x_2 - x_4). \end{aligned} \quad (7.18)$$

The relevance of the Wick theorem now becomes apparent: After taking the vacuum expectation value $\langle 0 | \dots | 0 \rangle$, only the total contraction survives. Hence, the vacuum expectation value is given entirely in terms of Feynman propagators D_F .

The **proof** proceeds by induction: Note first that the case $n = 1$ and the case $n = 2$ has been proven in Sect. 7.1.

We now perform the step from n to $n+1$. Without loss of generality, the $(n+1)$ st field can be taken to have the largest time argument: $x^0 \geq x_i^0 \forall i$. We then have

$$T\varphi\varphi_1\cdots\varphi_n = \varphi T\varphi_1\cdots\varphi_n = \varphi : \varphi_1\cdots\varphi_n : + \varphi (\text{all contractions}), \quad (7.19)$$

where in the last equality we made use of the Wick theorem for n . The claim follows if we succeed in proving the following Lemma:

Lemma: Let $x^0 \geq x_i^0 \forall i$. Then

$$\begin{aligned} \varphi : \varphi_1\cdots\varphi_n : &= : \varphi\varphi_1\cdots\varphi_n : + : \overline{\varphi\varphi_1}\cdots\varphi_n : + : \overline{\varphi\varphi_1\varphi_2}\cdots\varphi_n : \\ &+ \dots + : \overline{\varphi\varphi_1\cdots\varphi_n} : \end{aligned} \quad (7.20)$$

The proof is straightforward. First write

$$\begin{aligned} \varphi : \varphi_1\cdots\varphi_n : &= \varphi^c : \varphi_1\cdots\varphi_n : + \varphi^a : \varphi_1\cdots\varphi_n : \\ &= \varphi^c : \varphi_1\cdots\varphi_n : + : \varphi_1\cdots\varphi_n : \varphi^a + [\varphi^a, : \varphi_1\cdots\varphi_n :] \\ &=: \varphi\varphi_1\cdots\varphi_n : + [\varphi^a, : \varphi_1\cdots\varphi_n :], \end{aligned} \quad (7.21)$$

Now we employ the derivation property of the commutator,

$$[A, B_1 \cdots B_n] = [A, B_1]B_2 \cdots B_n + B_1[A, B_2]B_3 \cdots B_n + \dots + B_1 \cdots B_{n-1}[A, B_n], \quad (7.22)$$

and the following simple identity:

$$[\varphi^a, \varphi_i] = [\varphi^a, \varphi_i^c] = \langle 0 | \varphi^a \varphi_i^c | 0 \rangle = \langle 0 | \varphi \varphi_i | 0 \rangle = \langle 0 | T \varphi \varphi_i | 0 \rangle = \overline{\varphi \varphi_i}. \quad (7.23)$$

Here in the penultimate equality we used $x^0 \geq x_i^0 \forall i \in \{1, \dots, n\}$. With this, the lemma follows immediately, which also completes our proof of the Wick theorem.

7.3 The Feynman propagator

We want to demonstrate that the Feynman propagator, as defined earlier, is given by the following integral:

$$D_F(x - y) = \int \frac{d^4 p}{(2\pi)^4} \frac{i}{p^2 - m^2 + i\epsilon} \exp\{-ip(x - y)\} \Big|_{\epsilon \rightarrow 0}. \quad (7.24)$$

To prove this, let us first assume $x^0 > y^0$ and perform the p^0 integration. For this purpose, it is convenient to write the denominator as:

$$(p^0 - (p^0)_1) \cdot (p^0 - (p^0)_2) \quad (7.25)$$

where $(p^0)_{1,2} = \pm \sqrt{\vec{p}^2 + m^2 - i\epsilon} = \pm (\sqrt{\vec{p}^2 + m^2 - i\epsilon'})$. We want to view the p^0 -integration as a contour integration in the complex p^0 -plane and close the integration-contour such that the integrand is suppressed on the added piece. Since for $p^0 \rightarrow -i\infty$, we have $-ip^0(x^0 - y^0) \rightarrow -\infty$, we close the contour in the lower half plane, cf. Fig. 10. Now the integral is performed by simply picking up the residue. Comparing the result with our previous expression for $\langle 0 | \varphi(x) \varphi(y) | 0 \rangle$ completes the demonstration. An analogous calculation can be done for $x^0 < y^0$, where one has to close the contour in the upper half plane.

7.4 Feynman rules

Feynman rules are prescriptions associating certain graphic elements with mathematical expressions. They allow for systematically drawing pictures (Feynman diagrams) and using them to write down formulae for terms in the perturbative expansion of correlation functions. We will develop this using examples:

Suppressing the ‘0’ for ‘vacuum’ for brevity, consider e.g.

$$\left\langle T \varphi_1 \varphi_2 \varphi_3 \varphi_4 \exp \left(i \int d^4 x \mathcal{L}_{int}(\varphi) \right) \right\rangle$$

and work it out order-by-order in λ . At $\mathcal{O}(\lambda^0)$ one finds:

$$\begin{aligned} \langle T \varphi_1 \cdots \varphi_4 \rangle &= \overbrace{\varphi_1 \varphi_2 \varphi_3 \varphi_4} + \overbrace{\varphi_1 \varphi_2 \varphi_3 \varphi_4} + \overbrace{\varphi_1 \varphi_2 \varphi_3 \varphi_4} \\ &= \begin{array}{c} 1 \\ \bullet \\ | \\ \bullet \\ 2 \end{array} \begin{array}{c} 3 \\ \bullet \\ | \\ \bullet \\ 4 \end{array} + \begin{array}{c} 1 \quad 3 \\ \bullet \text{---} \bullet \\ \bullet \text{---} \bullet \\ 2 \quad 4 \end{array} + \begin{array}{c} 1 \quad 3 \\ \bullet \quad \bullet \\ \diagdown \quad \diagup \\ \bullet \quad \bullet \\ 2 \quad 4 \end{array}, \end{aligned} \quad (7.26)$$

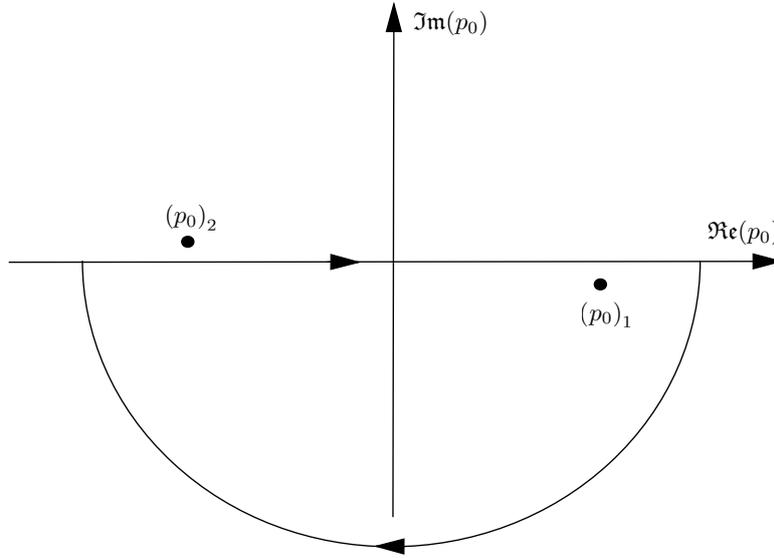


Figure 10: Pole structure and contour integration for Feynman propagator.

where $1 \bullet \text{---} \bullet 2 \equiv D_F(x_1, x_2)$.

Furthermore, at $\mathcal{O}(\lambda^1)$ one obtains

$$\begin{aligned}
 & \left\langle T\varphi_1 \cdots \varphi_4 \left(-\frac{i\lambda}{4!} \int d^4x \varphi(x)^4 \right) \right\rangle \\
 &= -\frac{i\lambda}{4!} \int d^4x \overline{\varphi}_1 \overline{\varphi}_2 \overline{\varphi}_3 \overline{\varphi}_4 \varphi \cdot 4! \\
 &+ \dots \text{terms in which } \overline{\varphi}\varphi \text{ appears once} \\
 &+ \dots \text{terms in which } \overline{\varphi}\varphi\overline{\varphi}\varphi \text{ appear} \\
 &= \sim \underbrace{\begin{array}{c} 1 \quad 3 \\ \diagdown \quad / \\ \bullet \quad \bullet \\ / \quad \diagdown \\ 2 \quad 4 \end{array}}_{\text{1st line}} + \sim \underbrace{\begin{array}{c} 1 \quad 3 \\ \text{---} \\ 2 \quad 4 \end{array} + \begin{array}{c} 1 \quad 3 \\ \text{---} \\ \text{loop} \\ \text{---} \\ 2 \quad 4 \end{array} + \sim \begin{array}{c} 1 \quad 3 \\ \text{---} \\ \text{loop} \\ \text{---} \\ 2 \quad 4 \end{array} + \dots}_{\text{2nd line}} \quad (7.27) \\
 &+ \sim \underbrace{\begin{array}{c} \text{loop} \end{array}}_{\text{3rd line}} \cdot (\text{Results of 7.26}) .
 \end{aligned}$$

Note that we have put proportionality signs in front of the diagrams since, as will become clear in a moment, there are non-trivial numerical prefactors with which the mathematical expressions for each diagram have to be weighted.

Let us now continue at $\mathcal{O}(\lambda^2)$:

$$\begin{aligned}
& \left\langle T \varphi_1 \cdots \varphi_4 \frac{1}{2!} \left(-\frac{i\lambda}{4!} \int d^4x \varphi(x)^4 \right) \left(-\frac{i\lambda}{4!} \int d^4y \varphi(y)^4 \right) \right\rangle \\
&= \frac{1}{2!} \left(-\frac{i\lambda}{4!} \right)^2 \int d^4x \int d^4y \overbrace{\varphi_1 \varphi_x \varphi_2 \varphi_x \varphi_x \varphi_y \varphi_x \varphi_y \varphi_y \varphi_3 \varphi_y \varphi_4} \cdot \# \\
& \quad + \text{other full contractions that give other Feynman diagrams} \tag{7.28} \\
&= \begin{array}{c} 1 \\ \bullet \\ \diagdown \\ x \bullet \circlearrowleft \bullet y \\ \diagup \\ 2 \end{array} \begin{array}{c} 3 \\ \bullet \\ \diagup \\ y \bullet \circlearrowleft \bullet x \\ \diagdown \\ 4 \end{array} + \dots
\end{aligned}$$

Here $\#$ is a number, contributing to the numerical prefactor mentioned above. The naive expectation is: One factor $4!$ from reshuffling φ_x , another factor $4!$ from reshuffling φ_y and finally $2!$ from $\varphi_x \leftrightarrow \varphi_y$. But the result of the explicit calculation (obtained by explicitly making all different contractions and adding them) is actually slightly different in this particular case. Note that the dots stand for all other Feynman diagrams (i.e. truly different pictures) which can be built from 6 propagators and two vertices. Examples are:

$$\begin{array}{c} \dots + \end{array} \begin{array}{c} 1 \\ \bullet \\ \diagdown \\ \bullet \circlearrowleft \bullet \\ \diagup \\ 2 \end{array} \begin{array}{c} 3 \\ \bullet \\ \diagup \\ \bullet \circlearrowleft \bullet \\ \diagdown \\ 4 \end{array} + \begin{array}{c} 1 \\ \bullet \\ \diagdown \\ \bullet \circlearrowright \bullet \\ \diagup \\ 2 \end{array} \begin{array}{c} 3 \\ \bullet \\ \diagup \\ \bullet \circlearrowright \bullet \\ \diagdown \\ 4 \end{array} + \dots \tag{7.29}$$

Having considered all these examples, it is straightforward to write down the general prescriptions for drawing Feynman diagrams:

- Each end of each propagator attaches either to an external point or to a vertex.
- Each external point accepts one and each vertex accepts four ends of a propagator.

Together with these prescriptions come the explicit rules for associating mathematical expressions and the elementary building blocks (lines and vertices) for our Feynman diagrams. They are called **Feynman rules** read, in $\lambda\varphi^4$ theory:

$$x \bullet \text{---} \bullet y = D_F(x - y) \tag{7.30}$$

$$\begin{array}{c} \diagdown \\ \bullet \\ \diagup \end{array} = (-i\lambda) \int d^4x .$$

Now for the numerical prefactors: Generally, the prefactor will be unity. In other words, by drawing a Feynman diagram and blindly applying the Feynman rules

above one gets the correct mathematical expression for the corresponding term in the λ -expansion of the correlation function in question.

The reason for this convenient result are our appropriately chosen conventions: The $1/4!$'s coming with φ^4 compensate the number of combinatorial possibilities associated with exchanging the corresponding fields. Also, the factor $1/2$ coming with the free lagrangian was crucial in obtaining our formula for the Feynman propagator - without any extra prefactor. Finally, the possible exchanges of different terms \mathcal{L}_{int} arising at higher order in the Taylor expansion of the exponential function are compensated the $1/n!$'s of the Taylor series.

Unfortunately, many diagrams are non-generic, in the sense of having symmetries and hence non-trivial prefactors, so called **symmetry factors**. To see an example, consider the following Feynman diagram:

$$\begin{array}{c}
 1 \text{ --- } 3 \\
 \phantom{1 \text{ --- } 3} \\
 2 \text{ --- } \text{O} \text{ --- } 4
 \end{array}
 \quad . \tag{7.31}$$

Write down the fields $(\varphi_1\varphi_3\varphi_2\varphi\varphi_4\varphi\varphi)$, apply Wick's theorem and count the number possible contractions giving the desired picture. Multiply by the $1/4!$ coming with the vertex:

$$\underbrace{\overbrace{\varphi_1\varphi_3}}_1 \underbrace{\overbrace{\varphi_2\varphi}}_4 \underbrace{\overbrace{\varphi_4\varphi}}_3 \underbrace{\overbrace{\varphi\varphi}}_1 \Rightarrow \frac{4 \cdot 3}{4!} = \frac{1}{2}. \tag{7.32}$$

Thus, this diagram has a symmetry factor of $1/2$. One can see where the symmetry factor comes from in this particular case: Our diagram does not change if the two downward-pointing ends of the vertex are swapped. More generally, the symmetry factor is $1/N$ where N is the **order of the automorphism group** of the diagram. We also note that programs for creating diagrams and the corresponding mathematical expressions exist, an example being "FeynArts".

In summary, we have so far found:

$$\begin{aligned}
 \langle 0|T\varphi_1 \dots \varphi_n \cdot e^{iS_{int}}|0\rangle &= \sum_{\text{all contractions}} \varphi_1 \dots \varphi_n \cdot \exp \left\{ -i\lambda \int \varphi^4 \right\} \\
 &= \left\{ \begin{array}{l} \text{Sum over all Feynman diagrams} \\ \text{(including symmetry factors)} \end{array} \right\}
 \end{aligned} \tag{7.33}$$

In addition, for any diagram one can 'split off' the so called 'vacuum bubbles', i.e. diagrams without external lines. This is illustrated in the following four examples:

$$\begin{array}{ccc}
 1 \text{ --- } 3 & & 1 \text{ --- } 3 \\
 & \rightarrow & \\
 2 \text{ --- } 4 & & 2 \text{ --- } 4
 \end{array} \tag{7.34}$$

$$\begin{array}{ccc}
1 & \text{---} & 3 \\
& \circlearrowleft & \\
2 & \text{---} & 4
\end{array}
\rightarrow
\begin{array}{ccc}
1 & \text{---} & 3 \\
& \circlearrowleft & \\
2 & \text{---} & 4
\end{array}
\quad (7.35)$$

$$\begin{array}{ccc}
& & 1 & \text{---} & 3 \\
& & & & \\
& & 2 & \text{---} & 4
\end{array}
\rightarrow
\left(\begin{array}{c} \circlearrowleft \end{array} \right) \cdot \left(\begin{array}{ccc} 1 & \text{---} & 3 \\ & & \\ 2 & \text{---} & 4 \end{array} \right)
\quad (7.36)$$

$$\begin{array}{ccc}
\circlearrowleft & \circlearrowleft & \begin{array}{ccc} 1 & & 3 \\ & \times & \\ 2 & & 4 \end{array}
\end{array}
\rightarrow
\left(\begin{array}{cc} \circlearrowleft & \circlearrowleft \end{array} \right) \cdot \left(\begin{array}{ccc} 1 & & 3 \\ & \times & \\ 2 & & 4 \end{array} \right)
\quad (7.37)$$

There is a very plausible (and easy-to-prove, cf. [1]) claim coming with these pictures:

$$\left\{ \begin{array}{c} \text{sum over all} \\ \text{Feynman diagrams} \\ \text{(with certain external lines)} \end{array} \right\} = \left\{ \begin{array}{c} \text{sum over all} \\ \text{Feynman diagrams} \\ \text{without vacuum bubbles} \end{array} \right\} \cdot \{1 + \text{sum over all vacuum bubbles}\} \quad (7.38)$$

Clearly, also have

$$\langle T \exp(iS_{int}) \rangle = \{1 + \text{sum over all vacuum bubbles}\} \quad (7.39)$$

Thus we finally obtain

$${}_H \langle 0 | T \varphi_1^H \dots \varphi_n^H | 0 \rangle_H = \frac{\langle 0 | T \varphi_1 \dots \varphi_n e^{iS_{int}} | 0 \rangle}{\langle 0 | T e^{iS_{int}} | 0 \rangle} = \left\{ \begin{array}{c} \text{sum over all} \\ \text{Feynman diagrams} \\ \text{without vacuum bubbles} \end{array} \right\}. \quad (7.40)$$

Note that the vacuum on the l.h. side of the equality is the interacting vacuum of the Heisenberg picture. On the right, we only have free fields and the free vacuum.

7.5 Feynman Rules in Momentum Space

According to LSZ, scattering amplitudes are determined by residues of poles of Fourier-transformed, time-ordered correlation functions. We thus have to translate our Feynman rules to momentum space. With

$$G(x_1, \dots, x_n) \equiv \langle T \varphi_1^H \dots \varphi_n^H \rangle \quad (7.41)$$

let us define

$$\tilde{G}(p_1, \dots, p_n) \equiv \int d^4x_1 e^{-ip_1x_1} \dots \int d^4x_n e^{+ip_nx_n} G(x_1, \dots, x_n). \quad (7.42)$$

Here we have chosen a convention by which the minus sign in the exponent is associated to incoming particles while the plus sign in the exponent is associated to outgoing particles. This is consistent with our discussion of LSZ.

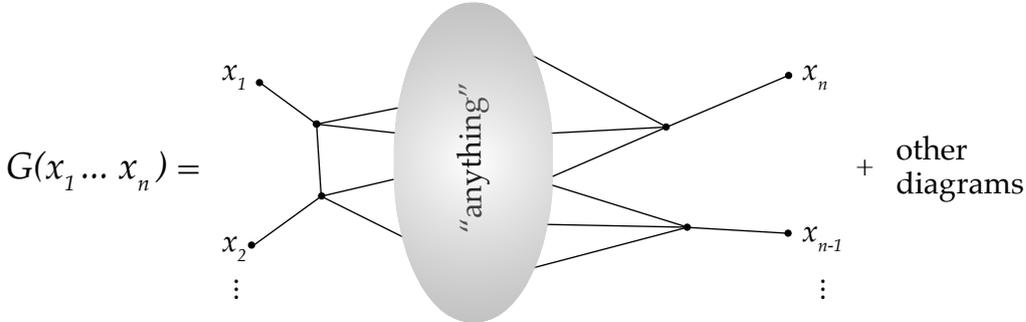
We note that the letter G is used because our time-ordered correlation functions are often also called **Green's functions**. This is justified since in particular the free 2-point function $G(x, y)$ is one of the inverses of the Klein-Gordon operator, cf. problems.

Recall that

$$x \bullet \text{---} \bullet y \equiv D_F(x - y) = \int \frac{d^4p}{(2\pi)^3} \frac{i}{p^2 - m_0^2 + i\varepsilon} e^{-ip(x-y)} \quad (7.43)$$

$$\begin{array}{c} \diagup \\ \bullet \\ \diagdown \end{array} \equiv -i\lambda \int d^4x$$

and consider some (in general complicated) real-space correlation function built from these two elements:



Clearly, in \tilde{G} the momentum-integration-variable p of every external line (cf. (7.43)) is fixed to the appropriate external momentum introduced with the Fourier transformation. This is due to the d^4x_i integration followed by d^4p integration.

Next, each d^4x integration of a vertex enforces momentum conservation at that vertex:

$$\begin{array}{c} p_1 \searrow \\ \diagup \\ \bullet \\ \diagdown \\ p_2 \nearrow \end{array} \begin{array}{c} \nearrow p_3 \\ \searrow p_4 \end{array} \propto -i\lambda \int d^4x e^{-ip_1x - ip_2x + ip_3x + ip_4x} \quad (7.44)$$

$$= -i\lambda (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - p_4).$$

Now, since each propagator end is either attached to a vertex or external, all $e^{\pm ipx}$ factors from the D_F 's are 'used up' by the last two calculational steps. They have disappeared from the overall expression.

Moreover, the momentum conservation δ functions at the vertices "kills" many of the d^4p integrations. But all those propagator momenta which are *not* fixed by the δ^4 -function from external or vertex d^4x integrations are still integrated over. The measure is $\int d^4p/(2\pi)^4$. They are called "loop momenta", for reasons to become clear momentarily in the following examples:

Example 1:

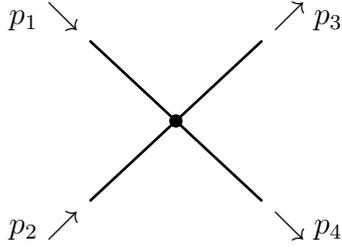
$$\begin{aligned}
p_1 \rightarrow \bullet \text{---} \bullet \rightarrow p_2 &= \int d^4x_1 e^{-ip_1x_1} \int d^4x_2 e^{+ip_2x_2} D_F(x_2 - x_1) \\
&= \frac{i}{p_1^2 - m_0^2 + i\varepsilon} (2\pi)^4 \delta^4(p_1 - p_2).
\end{aligned} \tag{7.45}$$

Example 2:

$$\begin{aligned}
p_1 \rightarrow \text{---} \bullet \text{---} \rightarrow p_2 & \begin{array}{c} \xrightarrow{q} \\ \circ \end{array} \\
&= \int d^4x_1 e^{-ip_1x_1} \int d^4x_2 e^{+ip_2x_2} \int d^4x D_F(x_2 - x) D_F(x - x_1) D_F(x - x) \\
&= \left(\frac{i}{p_1^2 - m_0^2 + i\varepsilon} \right)^2 (2\pi)^4 \delta^4(p_2 - p_1) (-i\lambda) \int \frac{d^4q}{(2\pi)^4} \frac{i}{q^2 - m_0^2 + i\varepsilon}
\end{aligned} \tag{7.46}$$

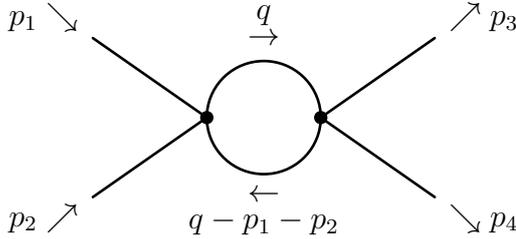
Two simple but key step in the calculation are $\delta^4(p_2 + q - p_1 - q) = \delta^4(p_2 - p_1)$ and $D_F(x - x) = \int \frac{i d^4q}{q^2 - m_0^2 + i\varepsilon}$. We see that a non-fixed propagator momentum does indeed arise because the diagram contains a loop.

Example 3:



$$\begin{aligned}
&= \int d^4x_1 e^{-ip_1x_1} \int d^4x_2 e^{-ip_2x_2} \int d^4x_3 e^{+ip_3x_3} \int d^4x_4 e^{+ip_4x_4} \\
&\quad \cdot (-i\lambda) \int d^4x D_F(x-x_1) D_F(x-x_2) D_F(x_3-x) D_F(x_4-x) \\
&= \left(\frac{i}{p_1^2 - m_0^2 + i\varepsilon} \right) \left(\frac{i}{p_2^2 - m_0^2 + i\varepsilon} \right) \left(\frac{i}{p_3^2 - m_0^2 + i\varepsilon} \right) \left(\frac{i}{p_4^2 - m_0^2 + i\varepsilon} \right) \\
&\quad \cdot (2\pi)^4 \delta^4(p_3 + p_4 - p_1 - p_2) (-i\lambda)
\end{aligned} \tag{7.47}$$

Example 4:



$$\begin{aligned}
&= \left(\frac{i}{p_1^2 - m_0^2 + i\varepsilon} \right) \left(\frac{i}{p_2^2 - m_0^2 + i\varepsilon} \right) \left(\frac{i}{p_3^2 - m_0^2 + i\varepsilon} \right) \left(\frac{i}{p_4^2 - m_0^2 + i\varepsilon} \right) \\
&\quad \cdot (2\pi)^4 \delta^4(p_3 + p_4 - p_1 - p_2) (-i\lambda)^2 \\
&\quad \cdot \underbrace{\int \frac{d^4q}{(2\pi)^4} \frac{i}{q^2 - m_0^2 + i\varepsilon} \frac{i}{(q - p_1 - p_2)^2 - m_0^2 + i\varepsilon}}_{\text{"loop integral"}} \cdot \{\text{symm. factor}\}
\end{aligned} \tag{7.48}$$

As already in Example 2, we again find a so-called loop integral as part of our expression – this time a slightly more complicated one.

From the examples above, we see that momentum-space correlation functions can, analogously to real-space correlation functions, be evaluated as sums of Feynman diagrams. The corresponding **momentum-space Feynman rules**

are:

$$\begin{array}{c} \bullet \text{---} \bullet \\ \xrightarrow{p} \end{array} = \frac{i}{p^2 - m_0^2 + i\varepsilon} \quad (7.49)$$

$$\begin{array}{c} \diagup \\ \bullet \\ \diagdown \end{array} = -i\lambda. \quad (7.50)$$

Additionally, the following prescriptions apply:

- Assign momenta at each vertex such that momentum conservation is ensured.
- Multiply by an overall factor of $(2\pi)^4 \delta^4(p_f - p_i)$.
- Multiply by $\int d^4p / (2\pi)^4$ for each closed loop, with p the momentum ‘running in the loop’.

7.6 Calculating the Z-Factor and the Physical Mass

Recall the spectral representation of our time-ordered correlation function of Heisenberg fields in the interacting vacuum:

$${}_H\langle 0|T\varphi^H(x)\varphi^H(y)|0\rangle_H = Z D_F(x-y, m^2) + \int_{M_t^2}^{\infty} d(M^2) \sigma(M^2) D_F(x-y, M^2). \quad (7.51)$$

We perform a Fourier-transform, drop the overall δ -function and give the result in the following diagrammatic representation:

$$p \bullet \text{---} \text{---} \bullet p = \frac{iZ}{p^2 - m^2 + i\varepsilon} + \int_{M_t^2}^{\infty} dM^2 \sigma(M^2) \frac{i}{p^2 - m^2 + i\varepsilon} \quad (7.52)$$

Here, by definition, $p \bullet \text{---} \text{---} \bullet p$ denotes all momentum-space Feynman diagrams

with two external lines, without vacuum bubbles and without an overall δ -function. Note that in expressions like (7.52), we may drop the $i\varepsilon$ since it is only relevant if we integrate p^0 and have to decide how to treat the poles on the integration contour.

Let us now write down the explicit Feynman-diagrams for the full 2-point func-

tion or **full propagator in momentum space** that we just defined:

$$\begin{aligned}
 & \text{Diagram: a circle with diagonal hatching and two external lines} \tag{7.53} \\
 = & \underbrace{\text{Diagram: a straight line with two external lines}}_{\lambda^0} + \underbrace{\text{Diagram: a straight line with a loop and two external lines}}_{\lambda^1} + \underbrace{\text{Diagram: a circle with two external lines}}_{\lambda^2} + \underbrace{\text{Diagram: a straight line with two loops and two external lines}}_{\lambda^2} + \dots
 \end{aligned}$$

Moreover, let us introduce the notation  for those diagrams which

do not fall apart upon cutting any internal line. With this definition, the previous series of diagrams can be written as

$$\text{Diagram: a circle with diagonal hatching and two external lines} = \text{Diagram: a straight line with two external lines} + \text{Diagram: a circle with cross-hatching and two external lines} + \text{Diagram: two circles with cross-hatching and two external lines} + \dots \tag{7.54}$$

We note that often it is also useful to consider certain diagrams or sums of diagrams *without* the external lines. For example, the diagrams which do not fall apart upon cutting an internal line, and with the external lines removed, may be denoted by double-shading and by dropping the dots which were used to denote the external

field positions. One then has . Such diagrams clearly do not fall apart

when any single line is cut. Hence, they are called **one-particle irreducible (1PI) diagrams**.

While the concept of 1PI diagrams is useful for any number of external lines (sometimes also call ‘legs’), here we are interested in the 1PI 2-point function. This is also known as the **self-energy** and denoted by $-i\Pi(p^2)$. From what has been said, it is clear that the following relation holds:

$$\text{Diagram: a circle with cross-hatching and two external lines} = \frac{i}{p^2 - m_0^2} (-i\Pi(p^2)) \frac{i}{p^2 - m_0^2} \tag{7.55}$$

It follows that

$$\begin{aligned}
\text{---} \bullet \text{---} \text{---} \text{---} \bullet \text{---} &= \frac{i}{p^2 - m_0^2} + \frac{i}{p^2 - m_0^2} (-i \Pi(p^2)) \frac{i}{p^2 - m_0^2} \\
&+ \frac{i}{p^2 - m_0^2} (-i \Pi(p^2)) \frac{i}{p^2 - m_0^2} (-i \Pi(p^2)) \frac{i}{p^2 - m_0^2} + \dots \quad (7.56) \\
&= \frac{i}{p^2 - m_0^2} \left(\frac{1}{1 - (-i \Pi(p^2))} \right) \left(\frac{i}{p^2 - m_0^2} \right) \\
&= \frac{i}{p^2 - m_0^2 - \Pi(p^2)}.
\end{aligned}$$

Combining this result with the spectral representation gives

$$\frac{i}{p^2 - m_0^2 - \Pi(p^2)} = \frac{iZ}{p^2 - m^2} + \underbrace{\int_{M_i^2}^{\infty} dM^2 \sigma(M)^2 \frac{i}{p^2 - M^2}}_{\text{no poles at } p^2 \approx m^2}. \quad (7.57)$$

The position of the pole and the residue have to match between the l.h. side and the first term on the r.h. side. Matching the poles implies

$$p^2 - m_0^2 - \Pi(p^2) = 0 \quad \text{at} \quad p^2 = m^2 \quad (7.58)$$

and hence

$$m^2 = m_0^2 + \Pi(m^2). \quad (7.59)$$

Matching residues means

$$\frac{p^2 - m^2}{p^2 - m_0^2 - \Pi(p^2)} \longrightarrow Z \quad \text{as} \quad p^2 \rightarrow m^2. \quad (7.60)$$

Taylor expanding $\Pi(p^2)$ at $p^2 = m^2$ and inverting the fraction gives

$$\frac{p^2 - m_0^2 - \{\Pi(m^2) + \Pi'(m^2)(p^2 - m^2)\}}{p^2 - m^2} \longrightarrow Z^{-1}. \quad (7.61)$$

Thus, we must have

$$Z^{-1} = 1 - \Pi'(m^2) \quad (7.62)$$

One sometimes calls this Z -factor the ‘wave function renormalization’, but this name will become clear only much later in our course.

7.7 Feynman rules for scattering amplitudes

The LSZ reduction formula may be written as

$$G(p_1, \dots, p_n, k_1, \dots, k_m) \sim \prod_j \frac{i\sqrt{Z}}{p_j^2 - m^2} \prod_i \frac{i\sqrt{Z}}{k_i^2 - m^2} \text{out} \langle p_1, \dots | k_1, \dots \rangle_{\text{in}}. \quad (7.63)$$

Moreover, in terms of Feynman diagrams the l.h. side takes the form

$$G(p_1, \dots, p_n, k_1, \dots, k_m) = \text{Diagram} = \text{Diagram} \quad (7.64)$$

where ‘A’ stands for **amputated diagram**, i.e. only diagrams which, after cutting any internal line, will not fall apart in such a way that precisely one external line is gone. By this definition, no external lines (i.e. propagators) are present in ‘A’. Note that ‘A’ is not exactly the same as 1 PI:

$$\text{Diagram} \neq \text{Diagram}$$

For example, the following diagram is a legitimate part of ‘A’:

$$\text{Diagram} \quad (7.65)$$

But is not part of the 1PI 4-point-function since it falls apart by cutting one line. Yet, it does not fall apart in such a way that *precisely one* external line is separated from the rest.

Furthermore we also know that:

$$\text{Diagram} \sim \frac{iZ}{p^2 - m^2} \quad (7.66)$$

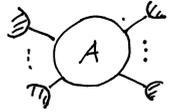
Using this relation and inserting (7.64) in (7.63), one can derive the following expression for the amplitude:

$$\text{out} \langle p_1, \dots, p_n | k_1, \dots, k_m \rangle_{\text{in}} = \left(Z^{\frac{1}{2}} \right)^{n+m} \cdot \left\{ \text{Diagram} \right\} \quad (7.67)$$

Equivalently we can formulate this result in terms of **Feynman rules** for $i\mathcal{M}_{\text{fi}}$:

The invariant amplitude $i\mathcal{M}_{\text{fi}}$ is calculated as the sum of all amputated diagrams without vacuum bubbles and without the overall δ -function. In addition, one has to multiply by a factor of $Z^{\frac{1}{2}}$ for each external line.

One may visualize the last statement by drawing the external lines as ‘’, i.e.

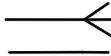
$$\text{out} \langle p_1, \dots, p_n | k_1, \dots, k_m \rangle_{\text{in}} = \text{} . \quad (7.68)$$

The other ‘half’ of the correction associated with the external line is absorbed in the normalization of the physical external particles.

We recall that our diagrams are built, in our specific theory, from the three ingredients

$$\text{} = \frac{i}{p^2 - m_0^2 + i\epsilon}; \quad \text{} = -i\lambda; \quad \int \frac{d^4p}{(2\pi)^4} \text{ for each loop.}$$

Sometimes, one also states that the diagrams used for the amplitude should be connected, i.e. that diagrams like

$$\text{} \quad (7.69)$$

are to be excluded. Indeed, this diagram is *not* part of 2-to-4 scattering since one momentum does not change. Thus, excluding the disconnected diagrams automatically follows if one applies the result to ask sensible physical questions.

Note that we could move forward and calculate cross-sections at loop level. But we would have trouble dealing with divergent diagrams. To solve this, one could perform a so called **Wick rotation** $p_0 \rightarrow ip_0$ and demand $p_E^2 < \Lambda^2$ for the euclidean momentum p_E . This is known as **cutoff regularization**. The subsequent limiting procedure of removing this cutoff (i.e. letting $\Lambda \rightarrow \infty$) is called **renormalization**. We will explain this important part of quantum field theory in the case quantum electrodynamics, where the divergences are weaker and the procedure is hence more intuitive.

8 The Electromagnetic Field

8.1 Gauge invariance

Recall the Lagrangian for the complex scalar:

$$\mathcal{L} = \partial_\mu \phi \partial^\mu \phi^* - m^2 \phi \phi^* . \quad (8.1)$$

We have already seen that it has a (global) U(1)-symmetry: $\phi(x) \rightarrow e^{i\alpha} \phi(x)$. In this context, ‘global’ means that this phase α is the same for every $x \in \mathbb{R}^{1,3}$.

Following our locality paradigm, we would like to promote this to a **local** or **gauge symmetry**:

$$\phi(x) \rightarrow e^{i\alpha(x)}\phi(x). \quad (8.2)$$

To figure out the transformation behaviour of \mathcal{L} , we analyse how the derivative of our scalar transforms:

$$\partial_\mu\phi \rightarrow \partial_\mu(e^{i\alpha}\phi) = e^{i\alpha}(\partial_\mu\phi + i(\partial_\mu\alpha)\phi). \quad (8.3)$$

This is *not* equal to $e^{i\alpha}\partial_\mu\phi$. Hence, unlike the global case, the phase does not drop out and \mathcal{L} is not invariant.

Thus, $\partial_\mu\phi$ does not transform just with a phase. But we would like to have a similar quantity that does. For this purpose, one introduces a **gauge connection** $A_\mu(x)$ and defines the **covariant derivative**:

$$D_\mu \equiv \partial_\mu + iA_\mu. \quad (8.4)$$

It transforms as follows:

$$D_\mu\phi \rightarrow D'_\mu\phi' = (\partial_\mu + iA'_\mu)e^{i\alpha}\phi = e^{i\alpha}(\partial_\mu\phi + i(\partial_\mu\alpha)\phi + iA'_\mu\phi). \quad (8.5)$$

For this to be equal to

$$e^{i\alpha}D_\mu\phi = e^{i\alpha}(\partial_\mu\phi + iA_\mu\phi), \quad (8.6)$$

we must demand

$$A'_\mu = A_\mu - \partial_\mu\alpha. \quad (8.7)$$

With this definition of how the vector field A_μ transforms under local phase rotations, $D_\mu\phi$ transforms with an overall phase factor. As a result,

$$\mathcal{L} = (D_\mu\phi)(D^\mu\phi)^* - m^2\phi\phi^* \quad (8.8)$$

is invariant.

Let us also give a slightly more formal and less pragmatic explanation of the concept of a gauge connection. This will help in building intuition:

Recall that a conventional derivative is defined as:

$$n^\mu\partial_\mu\phi = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon}(\phi(x + \epsilon n) - \phi(x)), \quad (8.9)$$

with n a unit vector in \mathbb{R}^4 (or any other \mathbb{R}^d).

In the presence of (local) gauge symmetry this does not make sense since the phases of $\phi(x)$ and $\phi(x + \epsilon n)$ are independent. Consequently, we cannot ‘compare’ these two quantities. If we want to know whether ϕ changes, which is the purpose of the derivative, we have to **parallel transport** ϕ from x to y

$$\phi(x) \rightarrow \underbrace{U(y, x)}_{\in U(1)}\phi(x) \quad (8.10)$$

in such a way that the (non-invariant and hence meaningless) difference of phase factors between x and y is removed. Our factor U above fulfils this purpose if

$$U'(y, x) = e^{i\alpha(y)}U(y, x)e^{-i\alpha(x)}. \quad (8.11)$$

As a result, $U(y, x)\phi(x)$ will transform just like a field at y and we can compare $\phi(x)$ and $\phi(y)$ in a sensible way:

$$n^\mu D_\mu \phi = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (\phi(x + \epsilon n) - U(x + \epsilon n, x)\phi(x)) . \quad (8.12)$$

If we assume U to be smooth and $U(x, x) = 1$, we have

$$U(x + \epsilon n, x) \equiv 1 - i\epsilon n^\mu A_\mu(x) + \dots , \quad (8.13)$$

where A_μ has been *defined* as the linear coefficient of the Taylor expansion of U . This definition of D_μ and the transformation properties of A_μ agree with our earlier definition.

Furthermore, given some $A_\mu(x)$, we can define

$$U(y, x, C) = \exp \left(i \int_{C(x, y)} A_\mu dx^\mu \right) . \quad (8.14)$$

This object, also known as a **Wilson line** depends on the path C connecting x and y . In our previous usage, we may think of $U(x + n\epsilon, x)$ as the phase associated with the straight line from x to $x + \epsilon n$. The name ‘gauge connection’ should now be more clear – it allows one to connect different points in a gauge-invariant manner.

Crucially, having introduced a new field A_μ , we *must* specify its dynamics, i.e. a gauge invariant action for the gauge connection itself.

To do so, we observe that the differential operator D_μ transforms as:

$$D_\mu \xrightarrow{\alpha} D'_\mu = e^{i\alpha} D_\mu e^{-i\alpha} . \quad (8.15)$$

We may check the equality of the two differential operators explicitly:

$$e^{i\alpha} (\partial_\mu + iA_\mu) e^{-i\alpha} = e^{i\alpha} (\partial_\mu e^{-i\alpha}) + \partial_\mu + iA_\mu = \partial_\mu + iA'_\mu \quad (8.16)$$

It immediately follows that the commutator transforms as

$$[D_\mu, D_\nu] \longrightarrow e^{i\alpha} [D_\mu, D_\nu] e^{-i\alpha} . \quad (8.17)$$

At the same time, we have

$$\begin{aligned} [D_\mu, D_\nu] &= \{\partial_\mu \partial_\nu + \partial_\mu iA_\nu + iA_\mu \partial_\nu - A_\mu A_\nu\} - \{\mu \leftrightarrow \nu\} \\ &= \{i(\partial_\mu A_\nu) + iA_\nu \partial_\mu + iA_\mu \partial_\nu\} - \{\mu \leftrightarrow \nu\} = iF_{\mu\nu} \end{aligned} \quad (8.18)$$

with

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu . \quad (8.19)$$

Here by $\{\mu \leftrightarrow \nu\}$ we mean the previous expression, but with μ and ν exchanged. Obviously, when subtracting such a term expressions symmetric in $\mu\nu$ are removed.

We now understand that, contrary to appearances, $[D_\mu, D_\nu]$ is *not* a differential operator. Indeed, all terms with ‘derivatives acting to the right’ have dropped out. As a result,

$$e^{i\alpha} [D_\mu, D_\nu] e^{-i\alpha} = e^{i\alpha} e^{-i\alpha} [D_\mu, D_\nu] = [D_\mu, D_\nu] , \quad (8.20)$$

in other words, the field strength tensor $F_{\mu\nu}$ is gauge invariant.

We can now propose the **scalar QED Lagrangian**:

$$\mathcal{L} = -\frac{1}{4e^2}F_{\mu\nu}F^{\mu\nu} + |D_\mu\phi|^2 - m^2|\phi|^2, \quad (8.21)$$

which is invariant under $\phi \rightarrow e^{i\alpha(x)}\phi$ and $A_\mu \rightarrow A_\mu - \partial_\mu\alpha$.

It is also Poincaré-invariant, with transformation rules

$$\begin{aligned} \text{Translation: } d^\mu : A'_\mu(x) &= A_\mu(x - d) \\ \text{Lorentz rotation: } \Lambda : A'_\mu(x) &= \Lambda_\mu{}^\nu A_\nu(\Lambda^{-1}x) \end{aligned} \quad (8.22)$$

We saw earlier that the vector field $\partial_\mu\varphi$ constructed from φ transforms in this way. Here, we declare A_μ to be a **fundamental vector field** and to have this property by definition.

A possibly more familiar form of the Lagrangian is obtained by the field redefinition $A_\mu \rightarrow eA_\mu$. More precisely, we introduce B_μ through the definition $A_\mu \equiv eB_\mu$ and subsequently rename B into A . The result is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + |D_\mu\phi|^2 - m^2|\phi|^2 \quad \text{with} \quad D_\mu = \partial_\mu + ieA_\mu. \quad (8.23)$$

In this form, it is more apparent that e is a coupling constant.

Before closing, let us make a (possibly advanced) comment: Electrodynamics or any other ‘abelian gauge theory’ has a much more elegant and natural definition in terms of differential forms. Namely, A is a 1-form,

$$A = A_\mu dx^\mu, \quad (8.24)$$

and the gauge parameter α is a 0-form. The latter is just another name for a scalar function. The the gauge-transformation reads

$$A \rightarrow A + d\alpha \quad (8.25)$$

and the field strength, which is a 2-form, is defined as

$$F = dA, \quad \text{with} \quad F = \frac{1}{2}F_{\mu\nu}dx^\mu \wedge dx^\nu. \quad (8.26)$$

The gauge field part of the action may be written as

$$\int \mathcal{L} = -\frac{1}{2} \int F \wedge \star F, \quad (8.27)$$

where the Hodge operator, mapping F to $\star F$, is defined at the component level by $(\star F)^{\mu\nu} = \epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma}/2$.

8.2 Gupta-Bleuler Quantization

We focus on the free theory first,

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \quad (8.28)$$

and attempt to quantize the system in the familiar way, treating each of the four components as an independent field:

$$\begin{aligned} \pi^\mu &= \frac{\partial \mathcal{L}}{\partial \dot{A}_\mu} = \frac{\partial}{\partial(\partial_0 A_\mu)} \left(-\frac{1}{4}F_{\rho\nu}F_{\sigma\tau}\eta^{\rho\sigma}\eta^{\nu\tau} \right) \\ &= -\frac{1}{2}F_{\rho\nu}\eta^{\rho\sigma}\eta^{\nu\tau} \frac{\partial}{\partial(\partial_0 A_\mu)} (\partial_\sigma A_\tau - \partial_\tau A_\sigma) \\ &= -\frac{1}{2}F_{\rho\nu} (\eta^{\rho 0}\eta^{\nu\mu} - \eta^{\rho\mu}\eta^{\nu 0}) = F^{\mu 0}. \end{aligned} \quad (8.29)$$

In particular we find $\pi^0 = 0$ as $F^{\mu\nu}$ is anti-symmetric. Situations like this represent a well-known problem of certain Hamiltonian systems and their quantization. For a deeper understanding the reader may want to consult Dirac's famous "Lectures on Quantum Mechanics" or Kugo's "Eichtheorie", or the book on gauge theories by Henneaux/Teitelboim, where the quantization of so-called **systems with constraints** is discussed.

In our course we will overcome this problem, which is characteristic of theories with gauge invariance, in a somewhat naive manner by **fixing the gauge**.

We choose the Lorenz⁸ gauge by demanding

$$\partial A \equiv \partial_\mu A^\mu = 0. \quad (8.30)$$

This allows us to use the modified Lagrangian

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{\lambda}{2}(\partial A)^2, \quad (8.31)$$

which gives the same equation of motion because ∂A appears quadratically. Thus, the variation of \mathcal{L} will still be linear in ∂A and hence vanish.

For simplicity, we set $\lambda = 1$, which is a special case of Lorenz or covariant gauges, known as **Feynman gauge**. The lagrangian then simplifies to

$$\begin{aligned} \mathcal{L} &= -\frac{1}{4} (2(\partial_\mu A_\nu)(\partial^\mu A^\nu) - \underbrace{2(\partial_\mu A_\nu)(\partial^\nu A^\mu)}_{\text{integrate by parts and drop total derivatives}}) - \frac{1}{2}(\partial A)^2 \\ &= -\frac{1}{2}(\partial_\mu A_\nu)(\partial^\mu A^\nu) + \frac{1}{2}(\partial A)^2 - \frac{1}{2}(\partial A)^2 + \text{tot. deriv.} \\ &= \frac{1}{2}(\partial_\mu A_\nu)(\partial^\mu A^\nu) - \frac{1}{2}(\partial A)^2 + \text{tot. deriv.}, \end{aligned} \quad (8.32)$$

which looks like a lagrangian for four massless, real scalars. However, the sign of the kinetic term of A_0 is wrong.

⁸Named after Ludvig Lorenz, not to be confused with the Hendrik Lorentz of Special Relativity.

By explicit calculation or by analogy to the real scalar case, we see that

$$\pi^\mu = -\dot{A}^\mu = (-\eta^{\mu\nu})\dot{A}_\nu, \quad (8.33)$$

which as expected gives the wrong sign for the A_0 canonical momentum. This will affect our quantization procedure.

We quantize by demanding:

$$\begin{aligned} [A, A] &= [\pi, \pi] = 0 \\ [A_\mu(\vec{x}), \pi^\nu(\vec{y})] &= i\eta_\mu{}^\nu \delta^{(3)}(\vec{x} - \vec{y}) = i\delta_\mu{}^\nu \delta^{(3)}(\vec{x} - \vec{y}). \end{aligned} \quad (8.34)$$

Now we may Fourier transform, introduce a and a^\dagger as linear combinations of the transformed fields, determine their commutation relations and go over to Heisenberg fields. This all works as before and one arrives at

$$\begin{aligned} A_\mu(x) &= \int d\tilde{k} \left(a_{\vec{k},\mu} e^{-ikx} + a_{\vec{k},\mu}^\dagger e^{ikx} \right) \\ [a, a] &= [a^\dagger, a^\dagger] = 0, \quad \left[a_{\vec{k},\mu}, a_{\vec{k}',\nu}^\dagger \right] = -\eta_{\mu\nu} 2k^0 (2\pi)^3 \delta^{(3)}(\vec{k} - \vec{k}'), \end{aligned} \quad (8.35)$$

where we have $k^0 = |\vec{k}|$ since $m = 0$. Note that again the sign of $[a_0, a_0^\dagger]$ is wrong.

The next step is to define $|0\rangle$ as the state annihilated by all $a_{\vec{k},\mu}, \forall \vec{k}, \mu$ and to construct the Fock space basis by applying all types of a^\dagger . However, this leads to two problems:

Problem 1: The wrong sign of the $[a_0, a_0^\dagger]$ renders our Hilbert space metric non-positive-definite. To see this, consider a harmonic oscillator and an algebra with the only non-trivial commutation relation $[a, a^\dagger] = -1$. For the first excited state one has

$$\|a^\dagger |0\rangle\|^2 = \langle 0| a a^\dagger |0\rangle = \langle 0| (a^\dagger a - 1) |0\rangle = -1, \quad (8.36)$$

which unacceptable in quantum mechanics. Switching the roles of a and a^\dagger does not allow us to solve this problem, since there is a relative sign change between $[a_0, a_0^\dagger]$ and $[a_i, a_i^\dagger]$ which is enforced by Lorentz-symmetry.

Problem 2: We cannot impose our gauge choice $\partial A = 0$ at the operator level since

$$[A_0, \partial A] = [A_0, \partial_0 A^0 + \partial_i A^i] = [A_0, \dot{A}_0] \neq 0, \quad (8.37)$$

where the commutator with the spatial derivatives vanishes because $[A_\mu(\vec{x}), A_\nu(\vec{y})] = 0$ for $\mu \neq \nu$. Obviously contradicts our gauge condition.

Gupta and Bleuler suggested the following resolution: Let F be the Fock space constructed above. and define $F_{\text{physical}} \subset F$ by

$$\partial A^a |\Psi\rangle = 0 \quad \Leftrightarrow \quad |\Psi\rangle \in F_{\text{physical}}. \quad (8.38)$$

Here the superscript a denotes the annihilator part of A . This definition implies that, for $|\Psi\rangle \in F_{\text{physical}}$,

$$\langle \Psi | \partial A | \Psi \rangle = \langle \Psi | [\partial A^a + \partial A^c] | \Psi \rangle = \langle \Psi | [\partial A^a + (\partial A^a)^\dagger] | \Psi \rangle = 0. \quad (8.39)$$

Here we used that $(\partial A)^\dagger$ vanishes acting to the left. As a result, our gauge condition is satisfied for expectation values on all physical states, which is sufficient for our purposes.

It will turn out that F_{physical} is positive-semi-definite, i.e. it contains no negative-norm-states. However, it does still have zero-norm states. We hence define

$$F_0 \equiv \{|\Psi\rangle \in F_{\text{physical}} : \|\Psi\rangle\| = 0\} \quad (8.40)$$

as the zero-norm subspace. This allows us to define our Hilbert space \mathcal{H} as

$$\mathcal{H} = F_{\text{physical}}/F_0. \quad (8.41)$$

This notation means that \mathcal{H} is the space of equivalence classes of vectors from F_{physical} , with the equivalence relation \sim defined by

$$|\Psi\rangle \sim |\Psi'\rangle \Leftrightarrow \|\Psi\rangle - |\Psi'\rangle\| = 0. \quad (8.42)$$

To see how all of this works more explicitly, it is convenient to use polarization vectors. Thus, a small interlude on polarization follows:

A general 1-photon state is a linear combination of states $a_{\vec{k},\mu}^\dagger |0\rangle$ with $\mu = 0, \dots, 3$. It may be written as

$$- \varepsilon^\mu(k) a_{\vec{k},\mu}^\dagger |0\rangle, \quad (8.43)$$

where \vec{k} is fixed and the minus sign has been introduced to avoid a minus from the metric when one considers space-like polarizations with lower index.

Now there are four independent polarizations for any given k and many possible basis choices. It is convenient to demand covariant orthonormality:

$$\varepsilon_\mu^{(\lambda)}(k) \overline{\varepsilon^{(\lambda')\mu}(k)} = \eta^{\lambda\lambda'}. \quad (8.44)$$

In the problems you are asked to demonstrate that this implies completeness, i.e.

$$\sum_{\lambda\lambda'} \eta_{\lambda\lambda'} \varepsilon_\mu^{(\lambda)}(k) \overline{\varepsilon_\nu^{(\lambda')}(k)} = \eta_{\mu\nu}. \quad (8.45)$$

To be fully explicit, we introduce some arbitrary but fixed unit vector $n = \{n^\mu\}$ with $n^2 = 1, n_0 > 0$ and demand

$$\varepsilon^{(0)} = n \quad \varepsilon^{(i)} \cdot n = 0 \quad \varepsilon^{(1)} \cdot k = \varepsilon^{(2)} \cdot k = 0. \quad (8.46)$$

The consistency of this requirement is most easily seen by going to the coordinate system in which (recall that $k^2 = m^2 = 0$)

$$n = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad k = |\vec{k}| \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \quad (8.47)$$

In this coordinate system, one easily shows that all of the above conditions are met by

$$\varepsilon^{(0)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \varepsilon^{(1)} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \varepsilon^{(2)} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \varepsilon^{(3)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \quad (8.48)$$

In fact, these ε 's are unambiguous up to rotations in the x - y -plane.

As a different common choice, which will be most useful for us, is based on a light-like auxiliary vector: We choose an $n \not\parallel k$ with $n^2 = 0$ and demand

$$\varepsilon^u = n, \quad \varepsilon^L \sim k, \quad \varepsilon^{(1)}, \varepsilon^{(2)} \text{ orthogonal to } n, k. \quad (8.49)$$

These ε 's form a basis but do not obey orthonormality. The previous orthonormality relations are replaced by

$$\begin{aligned} (\varepsilon^u)^2 &= (\varepsilon^L)^2 = 0 \\ \varepsilon^u \cdot \varepsilon^{(i)} &= \varepsilon^L \cdot \varepsilon^{(i)} = 0 \\ \varepsilon^{(i)} \cdot \varepsilon^{(j)} &= -\delta_{ij} \\ \varepsilon^u \cdot \varepsilon^L &= 1, \end{aligned} \quad (8.50)$$

where the second and third relation are the same as before. If we now go to a coordinate system such that

$$n = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \quad k = |\vec{k}| \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad (8.51)$$

we may use the explicit basis

$$\varepsilon^u = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \quad \varepsilon^L = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \varepsilon^{(1)} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \varepsilon^{(2)} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}. \quad (8.52)$$

Finally we define general 1-particle states by

$$-\varepsilon^\mu(k) a_{\vec{k}, \mu}^\dagger |0\rangle = |\varepsilon, k\rangle, \quad (8.53)$$

and obtain the scalar product

$$\begin{aligned} \langle \varepsilon', k' | \varepsilon, k \rangle &= \overline{\varepsilon'^\mu(k')} \varepsilon^\nu(k) \langle 0 | a_{\vec{k}', \mu} a_{\vec{k}, \nu}^\dagger | 0 \rangle \\ &= \overline{\varepsilon'^\mu(k')} \varepsilon^\nu(k) \left(-\eta_{\mu\nu} 2k^0 (2\pi)^3 \delta^{(3)}(\vec{k} - \vec{k}') \right) \\ &= -(\varepsilon' \cdot \varepsilon) 2k^0 (2\pi)^3 \delta^{(3)}(\vec{k} - \vec{k}'). \end{aligned} \quad (8.54)$$

Thus, $(-\varepsilon' \cdot \varepsilon)$ measures the overlap of states, and $-\varepsilon^2$ measures the norm of a single state. With this we are now ready to continue our discussion.

First, we reformulate our physical-state condition $\partial A^a |\Psi\rangle = 0$ in terms of polarizations:

$$\begin{aligned} \partial A^a |\varepsilon, q\rangle = 0 &\Leftrightarrow k_\mu a_{\vec{k}}^\mu |\varepsilon, q\rangle = 0 \\ \Leftrightarrow k_\mu a_{\vec{k}}^\mu a_{\vec{q}}^{\nu\dagger} \varepsilon_\nu(q) |0\rangle &\Leftrightarrow k \cdot \varepsilon(k) = 0. \end{aligned} \quad (8.55)$$

This condition is violated only for ε^u , which we thus call ‘‘unphysical’’ (explaining also our index choice).

Let us perform a linear transformation on the space of creation/annihilation operators, defining

$$\alpha_{\vec{k},(u,L,1,2)}^\dagger \equiv \varepsilon^{(u,L,1,2)\mu}(k) a_{\vec{k},\mu}^\dagger \quad (8.56)$$

We may now think of F as being built by applying the four α^\dagger 's to $|0\rangle$. This is of course still the same F as before.

As an important fact, we observe that F_{physical} is the subspace of F built by using only $\alpha_{\vec{k},(L,1,2)}^\dagger$. This clearly holds since

$$(q \cdot a_{\vec{q}}) \left(\text{products of various } \varepsilon_{\vec{k}}^{(L,1,2)} \cdot a_{\vec{k}}^\dagger \right) |0\rangle = 0. \quad (8.57)$$

Thus, we know that any state in F_{physical} can be written as superposition of states of the form

$$|\psi\rangle = \left(\text{products of } \alpha_{(L,1,2)}^\dagger \right) |0\rangle. \quad (8.58)$$

The next important fact is that, for such states $|\psi\rangle$, we have $\| |\psi\rangle \| = 0$ if and only if at least one α_L^\dagger appears in this product (since $\varepsilon_L^2 = 0$). In this case one also has $|\psi\rangle \sim 0$.

Finally, we note that the presence of such zero-norm states does not affect expectation values of observables. Crucially, in the context of a gauge theory observables are the operator versions of gauge-invariant classical expressions. For such operators we claim that

$$\langle \psi' | O | \psi' \rangle = \langle \psi | O | \psi \rangle \quad \text{if} \quad |\psi'\rangle = |\psi\rangle + (\dots \alpha_L^\dagger \dots) |0\rangle. \quad (8.59)$$

We will only demonstrate our claim for a simple example:

$$\begin{aligned} H =: \int d^3x \left(\pi^\mu \dot{A}_\mu - \mathcal{L} \right) : &= \dots = \int d\tilde{k} k_0 \left(-a_{\vec{k},\mu}^\dagger a_{\vec{k}}^\mu \right) \\ &= \int d\tilde{k} k_0 \left(\sum_{i=1}^2 \alpha_{\vec{k},i}^\dagger \alpha_{\vec{k},i} - \left[\alpha_{\vec{k},u}^\dagger \alpha_{\vec{k},L} + \alpha_{\vec{k},L}^\dagger \alpha_{\vec{k},u} \right] \right). \end{aligned} \quad (8.60)$$

The expression in square brackets vanishes ‘inside’ any physical state, $\langle \psi | [\dots] | \psi \rangle = 0$. As a result, it can be dropped before calculating expectation values of H . Then no annihilation or creation operators with index L are present in H and, as a result, product states with α_L^\dagger excitations do not contribute to $\langle \psi | H | \psi \rangle$.

Let us give a short **summary** of our findings:

- States in Fock space F are sums of products of α^\dagger 's acting on the vacuum:

$$|\psi\rangle \in F \Rightarrow |\psi\rangle = \sum \left(\alpha_{\vec{k},i}^\dagger \alpha_{\vec{p},u}^\dagger \alpha_{\vec{q},L}^\dagger \dots \right) |0\rangle.$$

- The physical subspace contains only states without unphysical polarizations:
 $|\psi\rangle \in F_{\text{physical}} \Rightarrow$ No terms involving α_u^\dagger appear.
- The zero-norm subspace is built only from product states involving at least one longitudinal polarization:
 $|\psi\rangle \in F_0 \Rightarrow$ Each term in the sum involves at least one α_L^\dagger factor.
(We note that this definition implies that F_0 is a linear subspace).
- Finally, the Hilbert space is built from F_{physical} by modding our F_0 :
 $\mathcal{H} \equiv F_{\text{physical}}/F_0 \equiv F_{\text{physical}}/\sim$, where \sim is an equivalence relation stating that two states are equivalent if they differ only by a vector from F_0 , i.e. by terms with α_L^\dagger .

We note that the freedom of adding states from F_0 (recall ε_L is parallel to \vec{k}) corresponds to residual gauge freedom of the classical theory:

$$\begin{aligned} A_\mu &\rightarrow A_\mu + \partial_\mu \chi \\ \text{Fourierspace : } \tilde{A}_\mu &\rightarrow \tilde{A}_\mu + ik_\mu \tilde{\chi}. \end{aligned} \tag{8.61}$$

If k is lightlike (which does not have to be the case for a classical gauge transformation), then adding a term $\sim k_\mu$ in Fourier space does not destroy our original gauge choice because taking a further derivative removes the term: $k^2 = 0$. The new field still obeys $k^\mu \tilde{A}_\mu = 0$. This explains the term ‘residual’ – it is the gauge freedom which is still left after Lorenz gauge has been imposed. The quantum analogue of the above is clearly the addition of states involving longitudinal polarizations.

We should recall that F , F_{physical} and \mathcal{H} were defined abstractly, before a specific choice for n . Hence, they do not depend on n – only our basis choice for the explicit description needed n .

Finally, we should emphasize that gauge symmetries differ from usual, global symmetries in a much deeper way than simply because of their x -dependence. Namely, both in the classical and in the quantum theory, states differing by a gauge transformation are declared to be **physically identical**. This is very different from, e.g., a particle on a line. In this case, in spite of perfect shift symmetry, x is a legitimate observable and states with different positions are distinguishable in principle. By contrast, two gauge-equivalent field configurations describe the same physical reality. In our construction, this become particularly clear since we **mod out** F_0 from our physical Fock space. What is physically real is encoded in the Hilbert space vector and gauge transformations simply do not exist in this final setting.

8.3 Photon Propagator

It is straightforward to calculate the vacuum expectation value of two photon fields,

$$\langle 0| A_\mu(x) A_\nu(y) |0\rangle = \langle 0| \int d\tilde{k} d\tilde{k}' e^{-ikx + ik'y} a_{\vec{k},\mu} a_{\vec{k}',\nu}^\dagger |0\rangle = -\eta_{\mu\nu} \int d\tilde{k} e^{-ik(x-y)}, \tag{8.62}$$

where we used that in the left field only a and in the right field only a^\dagger is relevant. We then find

$$\begin{aligned}
\langle 0|TA_\mu(x)A_\nu(y)|0\rangle &= \Theta(x^0 - y^0)\langle 0|A_\mu(x)A_\nu(y)|0\rangle + \{x \leftrightarrow y\} \\
&= \Theta(x^0 - y^0)\left(-\eta_{\mu\nu}\int d\tilde{k}e^{-ik(x-y)}\right) + \{x \leftrightarrow y\} \\
&= -\eta_{\mu\nu}(\Theta(x^0 - y^0)\langle 0|\varphi(x)\varphi(y)|0\rangle + \{x \leftrightarrow y\}) \\
&= -\eta_{\mu\nu}\langle 0|T\varphi(x)\varphi(y)|0\rangle = -\eta_{\mu\nu}D_F(x-y, m^2=0).
\end{aligned} \tag{8.63}$$

Our whole quantization procedure and the above derivation of the propagator can also be performed in a general gauge, $\lambda \neq 1$. The result is

$$\langle 0|TA_\mu(x)A_\nu(y)|0\rangle = \int \frac{d^4k}{(2\pi)^4}(-i)\left(\frac{\eta_{\mu\nu}}{k^2 + i\varepsilon} + \frac{1-\lambda}{\lambda}\frac{k_\mu k_\nu}{(k^2 + i\varepsilon)^2}\right)e^{-ik(x-y)}. \tag{8.64}$$

Elements of this will be discussed in the problems, see also [2]. Deriving this propagator is also rather straightforward in the path integral approach, to be studied in the next term.

8.4 Feynman rules for scalar QED

As a warm-up, consider a model with N different real scalars:

$$\mathcal{L} = \sum_{i=1}^N \frac{1}{2}(\partial\varphi^i)^2 - m^2(\varphi^i)^2 - \frac{\lambda}{8}\left(\sum_{i=1}^N(\varphi^i)^2\right)^2. \tag{8.65}$$

We note that from the many possible quartic interaction terms we chose one respecting the $O(n)$ symmetry of the free lagrangian. The propagator is

$$\langle T\varphi_1^i\varphi_2^j\rangle = \overline{\varphi_1^i\varphi_2^j} = \delta^{ij}D_F(x_1 - x_2) = \text{i} \bullet \text{---} \bullet \text{j} \tag{8.66}$$

To derive the vertex, we consider the simplest non trivial four-point function

$$\left\langle T\varphi_1^i\varphi_2^j\varphi_3^k\varphi_4^l \int d^4x \left(-\frac{i\lambda}{8}\right) (\delta_{mn}\varphi_x^m\varphi_x^n) (\delta_{pq}\varphi_x^p\varphi_x^q) \right\rangle, \tag{8.67}$$

and we focus only on the fully connected part:

$$\begin{array}{ccc}
& \text{j} & \text{l} \\
& \diagdown & \diagup \\
& \bullet & \\
& \diagup & \diagdown \\
& \text{i} & \text{k}
\end{array} \tag{8.68}$$

This means that $\varphi_1, \varphi_2, \varphi_3$ and φ_4 must each be contracted with one of the φ_x . First, consider the contraction

$$\overline{\varphi_1^i\varphi_x^m}\overline{\varphi_2^j\varphi_x^n}\overline{\varphi_3^k\varphi_x^p}\overline{\varphi_4^l\varphi_x^q} = \delta^{ij}\delta^{kl}D_F(x_1 - x)D_F(x_2 - x)D_F(x_3 - x)D_F(x_4 - x). \tag{8.69}$$

From this we obtain a contribution of

$$-\frac{i\lambda}{8}\delta^{ij}\delta^{kl} \quad (8.70)$$

to the vertex. The exact same contribution arises by exchanging $\varphi_x^m \leftrightarrow \varphi_x^n$ or $\varphi_x^p \leftrightarrow \varphi_x^q$ or $\varphi_x^m \varphi_x^n \leftrightarrow \varphi_x^p \varphi_x^q$. Thus, after adding these contractions, the factor $1/8$ disappears.

One can think of the sum of all those terms as arising from ‘pairing up’ i with j and k with l . There are two more such pairings: $(ik)(jl)$ and $(il)(jk)$. Hence:

$$\begin{array}{ccc} j & & l \\ & \diagdown & / \\ & \bullet & \\ & / & \diagdown \\ i & & k \end{array} = -i\lambda (\delta^{ij}\delta^{kl} + \delta^{ik}\delta^{jl} + \delta^{il}\delta^{jk}) . \quad (8.71)$$

As a cross check, note that $8 \cdot 3 = 4!$, so we did not forget anything.

As a side comment, we note that a maybe more elegant method to derive the vertex is to rewrite the interaction term as follows (see also problems):

$$-\frac{\lambda}{8} \left(\sum_{i=1}^N (\varphi^i)^2 \right)^2 = -\frac{\lambda}{4!} \sum_{ijkl=1}^N (\delta^{ij}\delta^{kl} + \delta^{ik}\delta^{jl} + \delta^{il}\delta^{jk}) \varphi^i \varphi^j \varphi^k \varphi^l . \quad (8.72)$$

Crucially, the combination of δ functions on the right provides a totally symmetric expression with four indices. Thus, we may treat this like our simplest example of a single real scalar, just with that scalar replaced by real vector full of scalars. But all the algebra, in particular the factor of $4!$ from the different contractions, remains unchanged and we may hence read off the vertex Feynman rule directly from the interaction term – we only need to add an i from $\exp(iS)$ and remove the $4!$.

As a further warmup consider a single complex scalar. Recall that

$$\phi(x) = \int d\tilde{k} \left(a_{\tilde{k}}^\dagger e^{ikx} + b_{\tilde{k}} e^{-ikx} \right) , \quad (8.73)$$

which implies

$$\overline{\phi_x \phi_y} = 0, \quad \overline{\phi_x^\dagger \phi_y} = \overline{\phi_x} \overline{\phi_y^\dagger} = D_F(x-y) . \quad (8.74)$$

Since ϕ and ϕ^\dagger are truly different, we can (even though D_F is symmetric in x and y) assign a direction to the corresponding line in the Feynman rule, for example

$$\overline{\phi_x \phi_y^\dagger} = y \longrightarrow x . \quad (8.75)$$

Here the arrow gives the direction of a b-particle, being created at y and annihilated at x . Note that in operator products time grows ‘right-to-left’ but in diagrams ‘left-to-right’.

After these preliminaries we simply state the Feynman rules of scalar QED. We will then give a partial derivation, which will be completed in the tutorials. Thus, the **Feynman rules for scalar QED** are:

Let us nevertheless focus on the previous simpler diagram and calculate the amplitude for the process (8.81) (see also problems):

$$\langle 0 | b_{\vec{p}'} \left(i \int d^4x \mathcal{L}_{\text{int}} \right) b_{\vec{p}}^\dagger \left(-\epsilon^\mu(k) a_{\vec{k},\mu}^\dagger \right) | 0 \rangle = i \mathcal{M}_{\text{fi}} (2\pi)^4 \delta^4(p' - k - p). \quad (8.83)$$

Here we just need to consider the cubic part of the interaction Lagrangian \mathcal{L}_{int} , meaning the part of $|D_\mu\phi|^2$ which contains the three fields A^μ , ϕ^\dagger and ϕ :

$$-ieA^\nu\phi^\dagger\partial_\nu\phi + \text{h.c.} \quad (8.84)$$

We obtain the following contributions:

$$\begin{aligned} A^\nu &\rightarrow \int d\tilde{q} a_{\vec{q}}^\nu e^{-iqx} \text{ acting on } a_{\vec{k},\mu}^\dagger |0\rangle \\ &\Rightarrow a_{\vec{q}}^\nu a_{\vec{k},\mu}^\dagger = -\eta^\nu{}_\mu (2\pi)^3 2k_0 \delta^3(\vec{k} - \vec{q}) \\ \partial_\nu &\rightarrow -ip_\nu \\ \phi &\rightarrow \int d\tilde{q}' b_{\vec{q}'} e^{-iq'x} \text{ acting on } b_{\vec{p}}^\dagger |0\rangle, \end{aligned} \quad (8.85)$$

and similarly for ϕ^\dagger .

Collecting everything we find:

$$i\mathcal{M}_{\text{fi}} = -iep^\mu\epsilon_\mu(k) \underbrace{-iep'^\mu\epsilon_\mu(k)}_{\text{vertex}} = -ie(p'^\mu + p^\mu) \underbrace{\epsilon_\mu(k)}_{\text{incoming photon}}. \quad (8.86)$$

from 2nd term in (8.84) incoming photon

Let us end with an important comment: Since the interactions involve \dot{A} , the canonical momentum of A receives a contribution from \mathcal{L}_{int} . Thus, the relation $\mathcal{H}_{\text{int}} = -\mathcal{L}_{\text{int}}$ does not hold and since \mathcal{H}_{int} is the crucial quantity in perturbation theory, our derivation above is not correct. However, at the same time, it is too naive to assume that ∂_μ commutes with contractions. For example:

$$\langle \text{T} \partial_\mu \varphi_x \partial_\nu \varphi_y^\dagger \rangle = \frac{\partial}{\partial x^\mu} \frac{\partial}{\partial y^\nu} \langle \text{T} \varphi_x \varphi_y \rangle - i\eta_{\mu 0} \eta_{\nu 0} \delta^4(x - y). \quad (8.87)$$

This effect also corrects the Feynman rules precisely compensating the error we made by assuming $\mathcal{H}_{\text{int}} = -\mathcal{L}_{\text{int}}$. In fact, this *had* to be the case to ensure that the final result is Poincaré-invariant. In summary, our naive analysis gave the correct result. For details see please consult the section on scalar electrodynamics in Sect. 6.1.4 of [2].

9 Spinors

9.1 Fields and representations

We already know three types of fields with different transformation properties:

$$\begin{aligned} \varphi(x) &\rightarrow \varphi(\Lambda^{-1}x) \\ A^\mu(x) &\rightarrow \Lambda^\mu{}_\nu A^\nu(\Lambda^{-1}x) \\ F^{\mu\nu}(x) &\rightarrow \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma F^{\rho\sigma}(\Lambda^{-1}x), \end{aligned} \quad (9.1)$$

with the last one not being elementary. This has an obvious extension to tensors with more than two indices.

Now we want to think about this in a more abstract way. Let us define our field as a map

$$\mathbb{R}^4 \longrightarrow V ; \quad x \mapsto \{\phi^i(x)\}, \quad (9.2)$$

where V denotes a vector space and i labels the corresponding components of the field in some basis. We note that a Lorentz-transformation acts in two ways:

- (1) On the argument (i.e. on spacetime \mathbb{R}^4): it acts in a unique, standard manner.
- (2) On the field value (i.e. on V): it acts differently from field to field.

Thus, each of our different fields is characterized by a vector space V and a **representation** of $SO(1,3)$ on V .

Let us briefly remind ourselves of this key concept: For any group G a representation R is a map $G \xrightarrow{R} GL(V)$ (general linear transformations on V) such that

$$R(\mathbb{1}) = \mathbb{1} \quad \text{and} \quad R(gh) = R(g)R(h). \quad (9.3)$$

In other words, R is a group homomorphism from G to $GL(V)$.

In our examples above we have:

- scalar: $V = \mathbb{R}$ or \mathbb{C} ; $R(\Lambda) = 1$ (trivial transformation)
- vector: $V = \mathbb{R}^4$; $R(\Lambda) = \Lambda$ (fundamental representation)
- tensor: $V = \mathbb{R}^4 \otimes \mathbb{R}^4$; $R(\Lambda) = \Lambda \otimes \Lambda$ (antisymmetric tensor representation).

We note that $F^{\mu\nu}$ or, more precisely, $F^{\mu\nu}\hat{e}_\mu \otimes \hat{e}_\nu$ lives in the antisymmetric subspace of $\mathbb{R}^4 \otimes \mathbb{R}^4$. Of course, fields taking values in the symmetric part also exist, for example the metric fluctuations (or graviton field).

We also note that, if we want to fit the tensor (e.g. our $F^{\mu\nu}$) into the general $\{\phi^i\}$ -notation, then the index i runs over all pairs of distinct indices μ and ν , such that the index pair $\{\mu\nu\}$ can be identified with the single index $\{i\}$. Our representation has $(4 \times 4 - 4)/2 = 6$ dimensions, which sets the range of $\{i\}$ and corresponds to the number different index pairs $\{\mu\nu\}$. As an advanced comment, we note that this 6-dimensional representation of $SO(1,3)$ naturally induces a representation of the subgroup $SO(3)$ of rotations. But this representation is **reducible**, i.e. it is the sum of two smaller representations of $SO(3)$. The latter of course the electric and magnetic field. In mathematical notation, these statements are formulated as $SO(1,3) \supset SO(3)$; $\mathbf{6} = \mathbf{3} + \mathbf{3}$. Here, in the last equation, the boldface numbers stand for the dimensions of the relevant representations and, at the same time, for the representations themselves.

9.2 Remarks on Lie Groups & Lie Algebras

Lie groups are groups which are also manifolds, such that the group operation is a diffeomorphism. (If manifolds are not yet known, the reader might think of smooth subspaces of \mathbb{R}^N and the group operation being differentiable.) Prime examples are the group $O(n)$ of orthogonal matrices, the group $U(n)$ of unitary matrices, and the $Sp(n)$ of symplectic matrices.

With a Lie group G always comes a Lie algebra $Lie(G) \equiv \mathfrak{g}$. A Lie algebra is a vector space \mathfrak{g} with a bilinear, antisymmetric map: $\mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$, $(a, b) \mapsto [a, b]$, satisfying the Jacobi-identity:

$$[a, [b, c]] + [c, [a, b]] + [b, [c, a]] = 0. \quad (9.4)$$

We will only need matrix groups and matrix Lie algebras (although the reader should be aware that important Lie groups which do not fall into this class exist). For matrix Lie groups, the relation between group and algebra can be understood in an elementary way (here we will not give any proofs):

The exponential map ‘exp’ of matrices represents a diffeomorphism from a small neighborhood of $0 \in M(n)$ to a small neighborhood of $\mathbb{1} \in Gl(n)$. Here $M(n)$ denotes the vector space of all $n \times n$ matrices and $Gl(n)$ the corresponding invertible matrices. Thus, we have

$$\exp(0) = \mathbb{1} \quad \text{and} \quad \exp(a) = g \quad (\text{where, for } a \text{ near } 0, g \text{ is near } \mathbb{1}). \quad (9.5)$$

Given a matrix group G , we may define $Lie(G)$ as the linear subspace of $M(n)$ generated by $\exp^{-1}(O_{\mathbb{1}})$, where $O_{\mathbb{1}}$ is a neighborhood of $\mathbb{1} \in G \subset M(n)$ (cf. Fig. 11). In other words, $\exp(a) = g$ with $a \in Lie(G), g \in G$ maps (at least) a small patch of $Lie(G)$ near 0 to a small patch of G near $\mathbb{1}$. In fact, this map extends to the whole Lie algebra and the image is the whole identity component of the group. But we will not derive this.

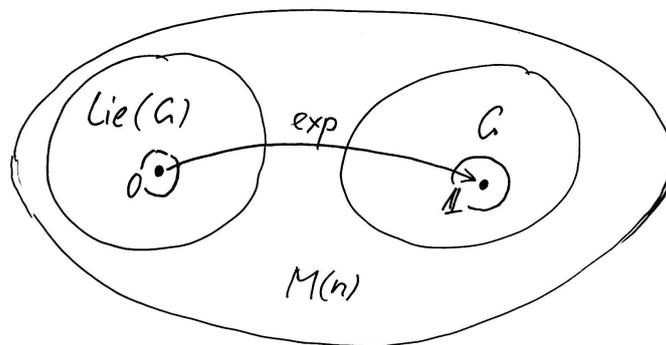


Figure 11: Illustration of group G and its Lie algebra as subsets of the set of all $n \times n$ matrices.

As an example, consider $G = SO(3)$. We claim that $Lie(G) = \{ \text{antisymmetric } 3 \times 3 \text{ matrices} \}$. Indeed, if $R = \exp(T)$ then

$$RR^T = \exp(T) \exp(T)^T = \exp(T) \exp(-T) = \mathbb{1} \quad (9.6)$$

for antisymmetric T . We note that the ‘ S ’ of $SO(n)$ is not visible at the Lie-algebra level.

It is illuminating to see in general that, if $a, b \in Lie(G)$, then $\exp [a, b] \in G$, where $[\cdot, \cdot]$ now is simply commutator rather than the abstract Lie algebra operation. Let $A = \exp(\epsilon a)$ and $B = \exp(\epsilon b)$, with ϵ denoting a small number. Clearly,

$$ABA^{-1}B^{-1} = C \in G. \quad (9.7)$$

We can expand this group element C as

$$\begin{aligned} C &= (\mathbb{1} + \epsilon a + \frac{\epsilon^2}{2} a^2)(\mathbb{1} + \epsilon b + \frac{\epsilon^2}{2} b^2)(\mathbb{1} - \epsilon a + \frac{\epsilon^2}{2} a^2)(\mathbb{1} - \epsilon b + \frac{\epsilon^2}{2} b^2) + \mathcal{O}(\epsilon^3) \\ &= \mathbb{1} + \epsilon^2 [a, b] + \mathcal{O}(\epsilon^3). \end{aligned} \quad (9.8)$$

Now we recall that $\exp(x)$ can be defined as the large- N -limit of $(1 + x/N)^N$. Thus, with $\epsilon^2 \equiv 1/N$ we find

$$\exp [a, b] = \lim_{N \rightarrow \infty} C^{1/N} \in G. \quad (9.9)$$

The point of this was to see, in a hands-on way, that it is really the matrix commutator which naturally allows one to build a new Lie algebra element from two given ones.

In analogy to groups, we also have the concept of a Lie algebra representation:

$$\text{Lie}(G) \xrightarrow{R} M(n), \quad a \mapsto R(a), \quad (9.10)$$

with $R(0) = 0$, $R([a, b]) = R(a)R(b) - R(b)R(a) = [R(a), R(b)]$.

In what follows, we will make extensive use of the following crucial fact which holds for simply connected⁹ Lie groups and their Lie algebras: **Given some representation R of a Lie algebra $\text{Lie}(G)$, we can always construct an associated representation of G** (which we will also call R by abuse of notation), such that

$$R(A) = \exp(R(a)) \quad \text{if} \quad A = \exp(a). \quad (9.11)$$

Let us sketch some of the key ideas of the proof: Define the desired group representation R by

$$R(A) \equiv \exp(R(\exp^{-1}(A))), \quad (9.12)$$

where for simplicity we assume that $A \in G$ is near the identity $\mathbb{1}$. All we need to show is that this R is fullfills

$$R(A) R(B) = R(A \cdot B). \quad (9.13)$$

To do so, we define $C \equiv A \cdot B$ and choose $a, b, c \in \text{Lie}(G)$ such that

$$A \equiv e^a, \quad B \equiv e^b, \quad C \equiv e^c. \quad (9.14)$$

We now need to show that

$$e^{R(a)} e^{R(b)} = e^{R(c)}. \quad (9.15)$$

Given our assumptions, we already know that

$$e^a e^b = e^c. \quad (9.16)$$

Moreover, the Baker-Campbell-Hausdorff formula tells us that

$$e^a e^b = e^{Z(a,b)}, \quad (9.17)$$

⁹ Roughly speaking, this means any loop in the group can be smoothly contracted to a point. We will later see why this qualification is important.

where

$$Z(a, b) = a + b + \frac{1}{2}[a, b] + \frac{1}{12}[a, [a, b]] - \frac{1}{12}[b, [a, b]] + \dots \quad (9.18)$$

is the famous series of iterated commutators that represents the heart of the Baker-Campbell-Hausdorff result. Hence,

$$c = Z(a, b). \quad (9.19)$$

This implies

$$e^{R(a)} e^{R(b)} = e^{Z(R(a), R(b))} = e^{R(Z(a, b))} = e^{R(c)}, \quad (9.20)$$

which is what we wanted to show. The key penultimate step relies in (9.20) relies on the fact R respects the commutator of the Lie algebra.

9.3 The spinor representation of $SO(1, 3)$

Let us first understand $Lie(SO(1, 3)) = \mathfrak{so}(1, 3)$. We write a Lorentz matrix that is close to the identity as an exponential and expand to leading order in the exponent:

$$\Lambda = e^{i\epsilon T} \quad \rightarrow \quad \Lambda_\mu^\nu = \delta_\mu^\nu + i\epsilon T_\mu^\nu + \dots \quad (9.21)$$

The i in the exponent is a conventional definition used by physicists - mathematicians would absorb it in T . The matrix T is by our definition an element of $\mathfrak{so}(1, 3)$ and ϵ is a small real parameter.

We recall

$$\Lambda_\mu^\nu \Lambda_\rho^\sigma \eta_{\nu\sigma} = \eta_{\mu\rho} \quad (9.22)$$

and investigate what this implies for T :

$$(\delta_\mu^\nu + i\epsilon T_\mu^\nu) (\delta_\rho^\sigma + i\epsilon T_\rho^\sigma) \eta_{\nu\sigma} = \eta_{\mu\rho} + \mathcal{O}(\epsilon^2) \quad \Rightarrow \quad T_{\mu\rho} + T_{\rho\mu} = 0. \quad (9.23)$$

In other words, after lowering the second index $SO(1, 3)$ generators are antisymmetric.

A useful basis of antisymmetric 4×4 matrices in the present context is one where each matrix $M_{\mu\nu}$ generates a (Lorentz) rotation in the μ - ν -plane. We define this basis by demanding that it fulfils the commutation relations

$$[M_{\mu\nu}, M_{\rho\sigma}] = i(\eta_{\nu\rho} M_{\mu\sigma} - \eta_{\mu\rho} M_{\nu\sigma} - \eta_{\nu\sigma} M_{\mu\rho} + \eta_{\mu\sigma} M_{\nu\rho}) \quad (9.24)$$

and leave it to the reader to find an explicit set of matrices satisfying this. The task is relatively easy since the basis deviates from the probably familiar basis of $Lie(SO(n))$ only by some signs.

A generic element of $Lie(SO(1, 3))$ can be written as

$$\Rightarrow \quad T_\mu^\nu = t^{\rho\sigma} (M_{\rho\sigma})_\mu^\nu. \quad (9.25)$$

We note that both $t^{\rho\sigma}$ and $M_{\rho\sigma}$ are antisymmetric, which implies that there are 6 independent basis elements and 6 parameters. Without providing a proof, we also note that *any* $\Lambda \in SO^+(1, 3)$ can be written as $\Lambda = \exp(it^{\mu\nu} M_{\mu\nu})$.

To define the spinor representation, we first introduce the **Clifford algebra** which is generated by $\mathbb{1}$ and four elements γ^μ which satisfy

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2\eta^{\mu\nu} \mathbb{1}, \quad (9.26)$$

where $\{\cdot, \cdot\}$ is the anti-commutator. Thus, the Clifford algebra is the vector space generated by $\mathbb{1}, \gamma^0, \gamma^1, \gamma^2, \gamma^3, \gamma^0\gamma^1, \gamma^1\gamma^2, \gamma^0\gamma^1\gamma^2, \dots$ with the relation $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}$ imposed. Two elements of this vector space can be multiplied to give another element - this makes it an algebra.

We will see that this algebra is finite-dimensional, a fact that is not immediately obvious. Much more of what follows could be done at this abstract level.

Nevertheless we want to use an explicit representation or realization of this algebra in terms of 4×4 matrices. One such representation is

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix}, \quad (9.27)$$

whereby every entry is 2×2 matrix and

$$\begin{aligned} \sigma^\mu &= (\sigma^0, \sigma^i) = \left\{ \mathbb{1}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\} \\ \bar{\sigma}^\mu &= (\sigma^0, -\sigma^i). \end{aligned} \quad (9.28)$$

Note that the overline here does not mean complex conjugation - the $\bar{\sigma}^\mu$ are simply a second set of 4 matrices. Furthermore we will use the notation

$$\gamma_\mu = \eta_{\mu\nu} \gamma^\nu, \quad (9.29)$$

and analogously for the σ 's.

We have to check whether those 4×4 matrices represent the Clifford algebra. To do so, consider

$$\{\gamma^\mu, \gamma^\nu\} = \left\{ \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix}, \begin{pmatrix} 0 & \sigma^\nu \\ \bar{\sigma}^\nu & 0 \end{pmatrix} \right\} = \begin{pmatrix} \sigma^\mu \bar{\sigma}^\nu + \sigma^\nu \bar{\sigma}^\mu & 0 \\ 0 & \bar{\sigma}^\mu \sigma^\nu + \bar{\sigma}^\nu \sigma^\mu \end{pmatrix}. \quad (9.30)$$

Now, analyze the cases $\mu, \nu = 0, 0 / 0, i / i, j$ separately, using also the probably familiar Pauli matrix relation $\{\sigma^i, \sigma^j\} = 2\delta^{i,j} \mathbb{1}$.

With this, the reader can straightforwardly check that the matrices

$$\mathcal{M}_{\mu\nu} \equiv \frac{i}{4} [\gamma_\mu, \gamma_\nu] \quad (9.31)$$

satisfy the same commutator relations as the $M_{\mu\nu}$ introduced earlier.

Thus the $\mathcal{M}_{\mu\nu}$ represent $\mathfrak{so}(1, 3)$ and we can construct a corresponding representation of $SO^+(1, 3)$.

The prescription is as follows: Write $\Lambda \in SO^+(1, 3)$ as

$$\Lambda = \exp(i t^{\mu\nu} M_{\mu\nu}). \quad (9.32)$$

Then define the action of Λ on \mathbb{C}^4 as:

$$\psi_D \xrightarrow{\Lambda} S(\Lambda) \cdot \psi_D \quad \text{where} \quad S(\Lambda) = \exp(i t^{\mu\nu} \mathcal{M}_{\mu\nu}). \quad (9.33)$$



Figure 12: Visualization of a double cover of a manifold.

This is a concrete realization of how we learned to abstractly lift a Lie algebra representation (here the one through the $\mathcal{M}_{\mu\nu}$) to a Lie group representation.

We are now ready for our main point: A **Dirac spinor** is a set of fields

$$(\psi_D)_a(x) \quad , \quad a = 1, 2, 3, 4 \quad (9.34)$$

transforming under the Lorentz group as

$$(\psi_D)_a(x) \xrightarrow{\Lambda} S(\Lambda)_a^b (\psi_D)_b(\Lambda^{-1}x). \quad (9.35)$$

Unfortunately, the full story is slightly more complicated. The reason is that our map $\Lambda \rightarrow S(\Lambda)$ is not defined globally on $SO^+(1, 3)$. To see this chose some arbitrary rotation axis and some corresponding generator $T = t^{\mu\nu}M_{\mu\nu}$. Let T be normalized such that a rotation around this axis is described by $\Lambda(\varphi) = \exp(i\varphi T)$. Naturally,

$$\Lambda(2\pi) = \mathbb{1}. \quad (9.36)$$

However, for $T_S = t^{\mu\nu}\mathcal{M}_{\mu\nu}$ one finds

$$S(2\pi) = \exp(i 2\pi T_S) = -\mathbb{1}. \quad (9.37)$$

This can be straightforwardly checked using our explicit realization of the Clifford algebra and hence of the generators. You have probably seen a similar situation arise in the transformation of spinor in quantum mechanics.

The issue is resolved as follows: The fundamental symmetry group of nature (ignoring translations for the moment) is *not* $SO(1, 3)$ but rather $\text{Spin}(1, 3)$. This is, by definition, the group generated by $\mathcal{M}_{\mu\nu}$'s. It is simply connected, so here we are really entitled to pass from any Lie algebra representation to a group representation. Specifically, the map $\mathcal{M}_{\mu\nu} \mapsto M_{\mu\nu}$ leads to an associated representation of this group acting on vectors:

$$\Lambda = \Lambda(S). \quad (9.38)$$

One may visualize what is going on by noting that $\text{Spin}(1, 3)$ is the “double cover” of $SO^+(1, 3)$, cf. Fig. 12. Note, however, that this visualization using S^1 as the double cover of another S^1 is not perfect. Both S^1 s are, of course, not simply connected. By contrast, the $2 : 1$ map from $\text{Spin}(1, 3)$ to $SO^+(1, 3)$ takes one from a simply connected to a non-simply-connected space.

The representation of $SO(1, 3)$ (more correctly $\text{Spin}(1, 3)$) on Dirac spinors is reducible. By this we mean that the relevant vector space is the direct sum of two

subspaces and the group acts independently on each of them. This becomes clear by explicitly working out the generators:

$$\mathcal{M}_{\mu\nu} = \frac{i}{4} [\gamma_\mu, \gamma_\nu] = \frac{i}{4} \left[\begin{pmatrix} 0 & \sigma_\mu \\ \bar{\sigma}_\mu & 0 \end{pmatrix}, \begin{pmatrix} 0 & \sigma_\nu \\ \bar{\sigma}_\nu & 0 \end{pmatrix} \right] = \begin{pmatrix} \sigma_\mu \bar{\sigma}_\nu - \sigma_\nu \bar{\sigma}_\mu & 0 \\ 0 & \bar{\sigma}_\mu \sigma_\nu - \bar{\sigma}_\nu \sigma_\mu \end{pmatrix}. \quad (9.39)$$

We note the block diagonal structure, which is of course retained after exponentiation. This confirms our claim, with the two relevant subspaces being formed by the two upper and the two lower components of our 4-component spinor respectively.

We thus can write

$$\psi_D = \begin{pmatrix} \psi_\alpha \\ \bar{\chi}^{\dot{\alpha}} \end{pmatrix}, \quad (9.40)$$

where the indices $\alpha, \dot{\alpha}$ run over 1, 2. Here the **Weyl spinor** ψ and the **complex-conjugate Weyl spinor** $\bar{\chi}$ transform independently. Note that, in addition to complex conjugation, the second Weyl spinor also has an upper rather than lower index. The meaning of this will be clarified shortly.

The decomposition of ψ_D in two independent parts can also be understood abstractly (i.e. without using our explicit representation of the γ 's). For this purpose, we first introduce

$$\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3 = \frac{i}{4!} \epsilon_{\mu\nu\rho\sigma} \gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma. \quad (9.41)$$

Now, due to the obvious property $\gamma^5 \gamma^\mu = -\gamma^\mu \gamma^5$, it follows that

$$\gamma^5 \mathcal{M}_{\mu\nu} = \mathcal{M}_{\mu\nu} \gamma^5. \quad (9.42)$$

Additionally one finds

$$(\gamma^5)^2 = \mathbb{1}. \quad (9.43)$$

With this, we may define

$$P_L \equiv \frac{1}{2} (\mathbb{1} - \gamma^5) \quad \text{and} \quad P_R \equiv \frac{1}{2} (\mathbb{1} + \gamma^5). \quad (9.44)$$

It follows immediately that $P_{L,R}$ satisfy

$$P_L^2 = P_L, \quad P_R^2 = P_R, \quad P_L + P_R = \mathbb{1}, \quad P_L P_R = 0. \quad (9.45)$$

The first two properties make P_L and P_R **projection operators**. The remaining two relations show that they induce a decomposition of the space on which they act in two subspaces:

$$V = V_L \oplus V_R \equiv \text{Im}(P_L) \oplus \text{Im}(P_R). \quad (9.46)$$

Since P_L and P_R commute with $\mathcal{M}_{\mu\nu}$, it now follows that $\psi_{D,L} \equiv P_L \psi_D$ and $\psi_{D,R} \equiv P_R \psi_D$ transform independently. This confirms what we already know from the explicit block-diagonal structure of the generators noted above. We will call $\psi_{D,L}$ and $\psi_{D,R}$ **left-handed** and **right-handed** Dirac spinors.

In our explicit representation we have

$$\gamma^5 = \begin{pmatrix} -\mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix}, \quad P_L = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & 0 \end{pmatrix}, \quad P_R = \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{1} \end{pmatrix}, \quad (9.47)$$

such that

$$\psi_{D,L} = \begin{pmatrix} \psi \\ 0 \end{pmatrix} \quad \text{and} \quad \psi_{D,R} = \begin{pmatrix} 0 \\ \bar{\chi} \end{pmatrix}. \quad (9.48)$$

Thus, the Weyl-spinor is a two-component object, containing the information of two complex numbers. Dirac spinors are four-component objects (vectors in \mathbb{C}^4), which are built from two Weyl spinors or, equivalently, from one left-handed and one right-handed Dirac spinor.

All of the above works for any even number of dimensions d . In this case the Dirac spinor has dimension $2^{d/2}$. For an odd number of dimensions, one uses $(d-1)$ -dimensional γ 's and adds $\gamma^d \propto \gamma^0 \gamma^1 \dots \gamma^{d-2}$. This γ^d does not commute with the $\mathcal{M}_{\mu\nu}$. Accordingly left- or right-handed Dirac as well as Weyl-spinors do not exist for an odd number of dimensions. (For a more detailed account see Polchinski, 'String Theory', vol. II, Appendix 'Spinors in various dimensions').

An interesting and useful fact special to $d = 4$ is that

$$\text{Spin}(1, 3) = \text{SL}(2, \mathbb{C}), \quad (9.49)$$

where $\text{SL}(2, \mathbb{C})$ is the group of 2×2 matrices M with $\det M = 1$.

Using this group isomorphism, the 2:1 map from $\text{Spin}(1, 3)$ to $\text{SO}(1, 3)$ can be given explicitly: Let $M \in \text{SL}(2, \mathbb{C})$, $v \in \mathbb{R}^4$ and $\hat{v} \equiv v_\mu \sigma^\mu$. Since $\{\sigma^\mu\}$ is a basis of hermitian 2×2 matrices, \hat{v} is a generic hermitian matrix.

Next we define $\hat{v}' = M \hat{v} M^\dagger$. Using this transformed matrix, we may implicitly define a transformed vector v' by $\hat{v}' = v'_\mu \sigma^\mu$. The straightforward calculation

$$\begin{aligned} (v')^2 &= (v'_0)^2 - (\vec{v}')^2 = \det \begin{pmatrix} v'_0 + v'_3 & v'_1 - i v'_2 \\ v'_1 + i v'_2 & v'_0 - v'_3 \end{pmatrix} \\ &= \det(\hat{v}') = \det(\hat{v}) = \det \begin{pmatrix} v_0 + v_3 & v_1 - i v_2 \\ v_1 + i v_2 & v_0 - v_3 \end{pmatrix} = v^2 \end{aligned} \quad (9.50)$$

demonstrates that the action of M on $\mathbb{R}^{1,3}$ which we have implicitly introduced preserves the length of vectors.

Let us repeat the logic: Any $M \in \text{SL}(2, \mathbb{C})$ defines a map $\hat{v} \mapsto \hat{v}' \equiv M \hat{v} M^\dagger$ on hermitian 2×2 matrices. Hence it also defines a map $v_\mu \mapsto v'_\mu$ that preserves the length. As a result, there exists a matrix $\Lambda = \Lambda(M) \in \text{SO}(1, 3)$ such that $v'_\mu = \Lambda_\mu{}^\nu v_\nu$. Obviously, $\Lambda(M) = \Lambda(-M)$, which is consistent with our claim that this map is 2:1.

We note without proof that our Weyl spinor ψ_α introduced earlier transforms as

$$\psi_\alpha \rightarrow M_\alpha{}^\beta \psi_\beta, \quad M \in \text{SL}(2, \mathbb{C}). \quad (9.51)$$

The other 2-spinor, which forms the two lower components of the 4-component Dirac spinor, is $\bar{\chi}^{\dot{\alpha}} = \epsilon^{\dot{\alpha}\dot{\beta}} \bar{\chi}_{\dot{\beta}}$, with $\bar{\chi}_{\dot{\beta}}$ being the complex-conjugate of a Weyl spinor: $\bar{\chi}_{\dot{\beta}} = (\chi_\beta)^*$ and $\dot{\beta} = \beta$. The complex-conjugate Weyl spinor transforms with the complex conjugate $SL(2, \mathbb{C})$ matrix:

$$\bar{\chi}_{\dot{\beta}} \rightarrow \bar{M}_{\dot{\beta}}{}^{\dot{\alpha}} \bar{\chi}_{\dot{\alpha}}. \quad (9.52)$$

This follows by complex conjugation from $\chi_\alpha \rightarrow M_\alpha^\beta \chi_\beta$. We also note that the 2-component ϵ -tensor is an invariant tensor of $SL(2, \mathbb{C})$, so it may naturally be used to lower and raise dotted and undotted Weyl indices.

It is a non-trivial claim that the Dirac spinor in our representation is formed precisely from a lower-index Weyl spinor and a complex-conjugate upper-index Weyl spinor. This claim could be checked explicitly using our definition of how $\psi_D = \begin{pmatrix} \psi \\ \bar{\chi} \end{pmatrix}$ transforms.¹⁰

To conclude, we could also say that $SL(2, \mathbb{C})$ is the fundamental symmetry group of space-time. Note that the relations between

$$SU(2) \subset SL(2, \mathbb{C}) \quad \text{and} \quad SO(3) \subset SO(1, 3) \quad (9.53)$$

and hence between spinors and vectors in non-relativistic quantum mechanics work very similarly.

9.4 Invariants involving spinors and lagrangian equations of motion

Our focus will now be on Dirac, not Weyl spinors. So we will drop the index D for brevity: $\psi_D \rightarrow \psi$. To write Lagrangians we need invariants, i.e. Lorentz-singlets built from ψ . Our first step in constructing such singlets is to recall that, for unitary representations of any symmetry group

$$v \rightarrow Uv, \quad U \in U(n), \quad (9.54)$$

the object $v^\dagger v \equiv \sum_i \bar{v}_i v_i$ is always invariant:

$$v'^\dagger v' = (Uv)^\dagger Uv = v^\dagger U^\dagger Uv = v^\dagger v. \quad (9.55)$$

This can also be written infinitesimally as

$$v'^\dagger v' \approx ((\mathbb{1} + iT)v)^\dagger (\mathbb{1} + iT)v = v^\dagger (\mathbb{1} - iT^\dagger) (\mathbb{1} + iT)v \approx v^\dagger (\mathbb{1} + i(T - T^\dagger))v. \quad (9.56)$$

Thus our claim follows from $T = T^\dagger$.

In our case ψ transforms as

$$\psi \rightarrow (\mathbb{1} + it^{\mu\nu} \mathcal{M}_{\mu\nu}) \psi. \quad (9.57)$$

From $(\gamma^0)^\dagger = \gamma^0$ and $(\gamma^i)^\dagger = -\gamma^i$ we can now conclude that

$$\mathcal{M}_{0i}^\dagger = -\mathcal{M}_{0i}, \quad \mathcal{M}_{ij}^\dagger = \mathcal{M}_{ij}. \quad (9.58)$$

Thus, $\psi^\dagger \psi$ is *not* an invariant. We therefore need to find a different object from which we can construct our Lagrangian. As a side-remark, we note that the mathematical

¹⁰ A standard reference for the very useful techniques of working with dotted and undotted Weyl indices is the famous book by Wess and Bagger [5] on supersymmetry (especially the Appendix). Some elements also appear in Chapter 1 of [6], in particular Problem 1.9.4.

reason for this is that $\text{SO}(1, 3)$ is non-compact. Non-compact Lie groups possess no finite-dimensional unitary representations.

To construct an invariant we observe that

$$\gamma^0 \gamma^\mu \gamma^0 = (\gamma^\mu)^\dagger, \quad \gamma^0 \mathcal{M}_{\mu\nu} \gamma^0 = \mathcal{M}_{\mu\nu}^\dagger, \quad \gamma^0 \mathcal{M}_{\mu\nu}^\dagger \gamma^0 = \mathcal{M}_{\mu\nu}. \quad (9.59)$$

It follows that $\psi^\dagger \gamma^0$ transforms according to

$$\psi^\dagger \gamma^0 \rightarrow \psi^\dagger (\mathbb{1} + it^{\mu\nu} \mathcal{M}_{\mu\nu})^\dagger \gamma^0 = \psi^\dagger (\mathbb{1} - it^{\mu\nu} \mathcal{M}_{\mu\nu}^\dagger) \gamma^0 = \psi^\dagger \gamma^0 (\mathbb{1} - it^{\mu\nu} \mathcal{M}_{\mu\nu}), \quad (9.60)$$

such that $\psi^\dagger \gamma^0 \psi$ is an invariant.

One generally defines

$$\bar{\psi} \equiv \psi^\dagger \gamma^0, \quad (9.61)$$

allowing for the very compact notation

$$\bar{\psi} \psi \quad (9.62)$$

for the invariant introduced above.

For the second ingredient in constructing Lorentz-singlets we use the following important relation, proved in one of the problems:

$$[\mathcal{M}_{\mu\nu}, \gamma_\rho] = -(\mathcal{M}_{\mu\nu})_\rho^\sigma \gamma_\sigma. \quad (9.63)$$

This basically says that, up to a sign, the commutator action of $\mathcal{M}_{\mu\nu}$ on the γ -matrices corresponds to the standard, multiplicative action of the $M_{\mu\nu}$. One may then conclude that

$$(\mathbb{1} + it^{\mu\nu} \mathcal{M}_{\mu\nu}) \gamma_\rho (\mathbb{1} - it^{\mu\nu} \mathcal{M}_{\mu\nu}) = (\mathbb{1} - it^{\mu\nu} M_{\mu\nu})_\rho^\sigma \gamma_\sigma + \mathcal{O}(t^2) \quad (9.64)$$

or, after exponentiation,

$$S(\Lambda) \gamma_\rho S(\Lambda)^{-1} = (\Lambda^{-1})_\rho^\sigma \gamma_\sigma. \quad (9.65)$$

If we now multiply by Λ from the left and make the spinor representation indices explicit, we find

$$\Lambda_\rho^\sigma (S(\Lambda))_a^b (\gamma_\sigma)_b^c (S(\Lambda)^{-1})_c^d = (\gamma_\rho)_a^d. \quad (9.66)$$

We have learned that $(\gamma_\sigma)_a^b$ is an **invariant tensor of $\text{SO}(1, 3)$** , where σ is a vector index, a is a Dirac spinor index, and b is an upper or inverse Dirac spinor index,¹¹ similar to the index of $\bar{\psi}^a$. Hence $\bar{\psi} \gamma^\mu \psi$ is a vector and $\bar{\psi} \gamma^\mu \psi v_\mu$ is a scalar.

For the Lagrangian we may thus write

$$\mathcal{L} = \bar{\psi} (i \gamma^\mu \partial_\mu - m) \psi. \quad (9.67)$$

This is the lowest order Lagrangian in fields and derivatives. In the problems it will be shown that the i is needed for S to be real.

¹¹ More formally, this is an index of the complex-conjugate Dirac spinor representation.

The equations of motion may now be straightforwardly derived: We treat $\psi, \bar{\psi}$ as independent variables and use the notation $\gamma^\mu v_\mu \equiv \not{v}$ (pronounced *v-slash*) for any Lorentz-vector or vector-operator v . It follows that¹²

$$0 = \delta S = \int d^4x \delta \mathcal{L} = \int d^4x \left[\delta \bar{\psi} (i\not{\partial} - m)\psi + \bar{\psi} (i\overleftarrow{\not{\partial}} - m)\delta\psi \right] \quad (9.68)$$

and hence

$$(i\not{\partial} - m)\psi = 0. \quad (9.69)$$

This is the celebrated **Dirac equation**.

As an important fact we observe that, if ψ solves the Dirac equation, then it also solves the Klein-Gordon-equation. To see this, we first note that, for any vector p , one has

$$\not{p}^2 = \gamma_\mu \gamma_\nu p^\mu p^\nu = \frac{1}{2} \{ \gamma_\mu, \gamma_\nu \} p^\mu p^\nu = \frac{1}{2} 2 \eta_{\mu\nu} p^\mu p^\nu = p^2. \quad (9.70)$$

By multiplying the Dirac equation with $(-i\not{\partial} - m)$ one then finds

$$0 = (-i\not{\partial} - m)(i\not{\partial} - m)\psi = (\not{\partial}^2 + m^2)\psi = (\partial^2 + m^2)\psi = 0, \quad (9.71)$$

confirming our claim.

9.5 Solutions of the Dirac equation

We make the ansatz $\psi(x) = u(p)e^{-ipx}$ with $p^2 = m^2$, justified by ψ solving the Klein-Gordon equation. We fix $p^0 > 0$, making our choice of sign in the exponent non-trivial. The other possible solution with e^{+ipx} will be considered below.

The Dirac equation constrains $u(p)$:

$$(i\not{\partial} - m)\psi = 0 \quad \Rightarrow \quad (\not{p} - m)u(p) = 0. \quad (9.72)$$

In a frame where $p = (m, \vec{0})$, this becomes $m(\gamma^0 - \mathbb{1})u(p) = 0$ or

$$\begin{pmatrix} -\mathbb{1} & \mathbb{1} \\ \mathbb{1} & -\mathbb{1} \end{pmatrix} u(p) = 0. \quad (9.73)$$

With the ansatz

$$u(p) = \begin{pmatrix} \xi \\ \xi' \end{pmatrix}, \quad (9.74)$$

we then have $\xi - \xi' = 0$. Since the space of 2-spinors is 2-dimensional, there are then two independent solutions. We write them as

$$u_s \sim \begin{pmatrix} \xi_s \\ \xi_s \end{pmatrix}, \quad s = 1, 2, \quad (9.75)$$

with $\xi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\xi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. It will be convenient to use the following normalization:

$$u_s(p) \equiv \sqrt{m} \begin{pmatrix} \xi_s \\ \xi_s \end{pmatrix} \quad \text{in a frame where } p = (m, \vec{0}). \quad (9.76)$$

¹²Note that $\overleftarrow{\partial}_\mu$ acts to the left and has an extra minus.

We note that, once defined in this particular frame, our solutions defined in any other frame by Lorentz transformations with $S(\Lambda)$.

Let us comment on the relation of Dirac spinors to non-relativistic spinors, which is particularly obvious in the frame $p = (m, \vec{0})$. To see this, let us exclude boosts, restricting attention to $SO(3) \subset SO(1, 3)$. At the Lie algebra level, this corresponds to the restriction

$$t^{\mu\nu} \mathcal{M}_{\mu\nu} \rightarrow t^{jk} \mathcal{M}_{jk} = t^{jk} \frac{i}{4} \begin{pmatrix} \sigma_j \bar{\sigma}_k - \sigma_k \bar{\sigma}_j & 0 \\ 0 & \bar{\sigma}_j \sigma_k - \bar{\sigma}_k \sigma_j \end{pmatrix}. \quad (9.77)$$

Using $[\sigma_j, \sigma_k] = 2i\varepsilon_{jkl}\sigma_l$, this becomes

$$t^{jk} \mathcal{M}_{jk} = t^{jk} \frac{1}{2} \begin{pmatrix} \varepsilon_{jkl}\sigma_l & 0 \\ 0 & \varepsilon_{jkl}\sigma_l \end{pmatrix}. \quad (9.78)$$

We see that both upper and lower two-component solutions transform under rotations just as in non-relativistic quantum mechanics.

Even more explicitly: To rotate around the 3-axis by an angle φ , we must pick

$$t^{jk} = \frac{1}{2} \varepsilon^{jkl} (\hat{e}_3)_l \varphi. \quad (9.79)$$

The reader should check that this is indeed the correct normalization using the explicit $SO(3)$ generators M_{ij} given earlier. Then, for the corresponding rotation of the spinor, one finds:

$$\Rightarrow \exp(it^{ij} \mathcal{M}_{ij}) = \begin{pmatrix} \exp(i\varphi \frac{1}{2} \sigma_3) & 0 \\ 0 & \exp(i\varphi \frac{1}{2} \sigma_3) \end{pmatrix}. \quad (9.80)$$

This is also consistent with $SU(2) \subset SL(2, \mathbb{C})$ and the $SL(2, \mathbb{C})$ -action on spinors described earlier. Moreover, appealing to what should be known from quantum mechanics, this demonstrates that our Dirac spinors are going to describe spin- $\frac{1}{2}$ -particles.

As advertised at the beginning, there is a second set of so-called **negative frequency** solutions: $\psi(x) = v(p)e^{+ipx}$ with $p^0 > 0, p^2 = m^2$. This time we have

$$(i\not{\partial} - m)\psi = 0 \quad \Rightarrow \quad (\not{p} + m)v(p) = 0. \quad (9.81)$$

In a frame with $p = (m, \vec{0})$ one has

$$\begin{pmatrix} \mathbb{1} & \mathbb{1} \\ \mathbb{1} & \mathbb{1} \end{pmatrix} v(p) = 0 \quad (9.82)$$

and hence

$$v_s(p) = \sqrt{m} \begin{pmatrix} \eta_s \\ -\eta_s \end{pmatrix}, \quad s = 1, 2 \quad (9.83)$$

with $\eta_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\eta_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Our choice of basis has been made to ensure the following relations:

$$\begin{aligned} \bar{u}_r(p) u_s(p) &= 2m \delta_{rs} & \bar{u}_r(p) v_s(p) &= 0 \\ \bar{v}_r(p) v_s(p) &= -2m \delta_{rs} & \bar{v}_r(p) u_s(p) &= 0. \end{aligned} \quad (9.84)$$

Crucially, by Lorentz invariance this holds in all frames since these expressions are Lorentz singlets.

In addition to these ‘orthonormality relations’, there is also a form of a ‘completeness relation’:

$$\begin{aligned} \sum_{s=1}^2 (u_s(p))_a (\bar{u}_s(p))^b &= (\not{p} + m)_a^b \\ \sum_{s=1}^2 v_s(p) \bar{v}_s(p) &= \not{p} - m. \end{aligned} \tag{9.85}$$

Here, in the second line, we used a very convenient shorthand, matrix notation in which the indices of the spinors are suppressed. The idea is that v_s is interpreted as a column-vector while \bar{v}_s is a row-vector. The product is then a matrix, which is equal to matrix on the r.h. side. Here m is shorthand for $m\mathbb{1}$.

We derive the first equation of the two relations in (9.85): For this purpose, let both sides of the equation act on the basis $\{u_s(p), v_r(p)\}$ of the spinor space \mathbb{C}^4 . Using the relations (9.84), we have:

$$\begin{aligned} \text{LHS} : \left(\sum_{s=1}^2 u_s(p) \bar{u}_s(p) \right) u_r(p) &= \sum_{s=1}^2 u_s(p) 2m \delta_{rs} = 2m u_r(p) \\ \left(\sum_{s=1}^2 u_s(p) \bar{u}_s(p) \right) v_r(p) &= 0 \end{aligned} \tag{9.86}$$

$$\begin{aligned} \text{RHS} : (\not{p} + m) u_r(p) &= \underbrace{(\not{p} - m) u_r(p)}_{=0} + 2m u_r(p) = 2m u_r(p) \\ (\not{p} + m) v_r(p) &= 0. \end{aligned} \tag{9.87}$$

Analogous manipulations demonstrate the validity of the second of the two relations in (9.85).

As a final comment, we note that it is easy to remember the signs in the completeness relations (9.85): Indeed, these relations must be consistent with the Dirac equation, which is the case only for the right sign choice:

$$(\not{p} - m) \sum_s u_s(p) \bar{u}_s(p) = (\not{p} - m)(\not{p} + m) = p^2 - m^2 = 0 \tag{9.88}$$

10 Quantization of Spinors

10.1 Hamiltonian

To transit from the Lagrangian to the Hamiltonian description, we first compute the canonical momenta for the set of 4 fields $\psi = \{\psi_a\}$. From the lagrangian density

$$\mathcal{L} = \bar{\psi}(i\not{\partial} - m)\psi \tag{10.1}$$

it follows that

$$\pi^a = \frac{\partial \mathcal{L}}{\partial \dot{\psi}_a} = \frac{\partial}{\partial \dot{\psi}_a} (i\psi^\dagger \gamma^0 \gamma^0 \dot{\psi}) = i(\psi^\dagger)^a. \tag{10.2}$$

We may write this in vector notation, with ψ a column vector and π a row vector as

$$\pi = i\psi^\dagger. \quad (10.3)$$

We note that the Lagrangian (and hence Hamiltonian) may be expressed using just ψ and $\pi = i\psi^\dagger$. There is hence no need for a canonical momentum corresponding to ψ^\dagger . This is different from the complex scalar case, where both ϕ, π and $\phi^\dagger, \pi^\dagger$ were needed.

The underlying reason for this distinction is the following: The equation of motion is only of first order in t in the case of the Dirac spinor. As a result, there is an effective reduction of the number of degrees of freedom (for more details, see Weinberg [4], Chapter 7).

Thus, we are now ready derive the Hamiltonian density in the standard way:

$$\begin{aligned} \mathcal{H} = \pi\dot{\psi} - \mathcal{L} &= i\psi^\dagger\dot{\psi} - \psi^\dagger\gamma^0(i\cancel{\partial} - m)\psi \\ &= -\psi^\dagger\gamma^0(i\gamma^i\partial_i - m)\psi \\ &= i\pi\gamma^0(i\gamma^i\partial_i - m)\psi. \end{aligned} \quad (10.4)$$

10.2 Quantization attempts with commutators

We could attempt to define, as we did for the scalar and vector case,

$$[\psi(\vec{x}), \pi(\vec{y})] = [\psi(\vec{x}), i\psi^\dagger(\vec{y})] = i\delta^3(\vec{x} - \vec{y})\mathbb{1}. \quad (10.5)$$

One would then proceed exactly as before. We skip these familiar intermediate steps and jump directly to the expression for the free fields in terms of creation and annihilation operators:

$$\psi(x) = \int d\tilde{p}(a_{\tilde{p}}^s u_s(p)e^{-ipx} + b_{\tilde{p}}^{s\dagger} v_s(p)e^{ipx}). \quad (10.6)$$

We would expect these operators to obey

$$\left[a_{\tilde{p}}^s, a_{\tilde{q}}^{r\dagger} \right] = (2\pi)^3 \delta^3(\tilde{p} - \tilde{q}) \delta^{rs} 2p^0 = \left[b_{\tilde{p}}^s, b_{\tilde{q}}^{r\dagger} \right]. \quad (10.7)$$

Of course, we must check the consistency with the original commutation relations (at $x^0 = y^0$):

$$\begin{aligned} [\psi(\vec{x}), \psi^\dagger(\vec{y})] &= \int d\tilde{p} d\tilde{q} \left(e^{i\tilde{p}\vec{x} - i\tilde{q}\vec{y}} u_s(p) u_r^\dagger(q) \left[a_{\tilde{p}}^s, a_{\tilde{q}}^{r\dagger} \right] + e^{-i\tilde{p}\vec{x} + i\tilde{q}\vec{y}} v_s(p) v_r^\dagger(q) \left[b_{\tilde{p}}^{s\dagger}, b_{\tilde{q}}^r \right] \right) \\ &= \int d\tilde{p} \left(e^{i\tilde{p}(\vec{x} - \vec{y})} (\cancel{p} + m) - \underbrace{e^{-i\tilde{p}(\vec{x} - \vec{y})} (\cancel{p} - m)}_{\tilde{p} \rightarrow -\tilde{p}} \right) \gamma^0 \\ &= \int d\tilde{p} e^{i\tilde{p}(\vec{x} - \vec{y})} (p_0 \gamma^0 + p_i \gamma^i + m - (p_0 \gamma^0 - p_i \gamma^i - m)) \gamma^0. \end{aligned} \quad (10.8)$$

Here, in the step from the first to the second line, we used the δ -function from the commutator and the relations $u_s(p) u_r^\dagger(p) = u_s(p) \bar{u}_r(p) \gamma^0 = (\cancel{p} + m) \gamma^0$ and similarly for the second term.

However, the result obtained in the last line is unsatisfactory: The p^0 term required to cancel the $1/p^0$ from $d\tilde{p}$ drops out. By contrast, the terms with mass and p^i do not cancel. We can not obtain the δ -function on the r.h. side of (10.5) in this way.

So let us try to assume instead that $[b, b^\dagger] = -1$, effectively exchanging the roles of b and b^\dagger . At first glance, this appears to work. Repeating the analysis one now finds

$$[\psi(\vec{x}), \psi^\dagger(\vec{y})] = \dots = \int d\tilde{p} e^{i\vec{p}(\vec{x}-\vec{y})} 2p^0 \gamma^0 \gamma^0 = \delta^3(\vec{x}-\vec{y}) \mathbb{1}. \quad (10.9)$$

However, proceeding to the quantum Hamiltonian in a straightforward manner reveals a deep problem. One finds

$$H = \int \mathcal{H} d^3x = \int d\tilde{p} p^0 \sum_s (a_{\vec{p}}^{s\dagger} a_{\vec{p}}^s - b_{\vec{p}}^{s\dagger} b_{\vec{p}}^s), \quad (10.10)$$

with a wrong sign in front of the antiparticle contribution. Due to the relative sign between the two terms, the energy is unbounded from below. In other words, the vacuum is unstable. No cure for this problem has been found in the framework which we have so far developed.

10.3 Quantization with anticommutators

The only known cure for the problem we just found is to fundamentally change of the quantization procedure. It can be proven rigorously that this is unavoidable for all fields with half-integer spin. This is known as the **Spin-Statistics Theorem**, see, e.g. [7], and it is related to the double-cover feature of the $Spin(1, 3)$ with respect to $SO(1, 3)$.

The key idea is to use anti-commutators, $\{a, b\} = ab+ba$, instead of commutators, when quantizing the theory. We basically repeat the previous subsection with the new postulate

$$\{\psi(\vec{x}), \pi(\vec{y})\} = \{\psi(\vec{x}), i\psi^\dagger(\vec{y})\} = i\delta^3(\vec{x}-\vec{y})\mathbb{1}. \quad (10.11)$$

It follows that the creation/annihilation operators obey

$$\left\{ a_{\vec{p}}^r, a_{\vec{q}}^{s\dagger} \right\} = \left\{ b_{\vec{p}}^r, b_{\vec{q}}^{s\dagger} \right\} = (2\pi)^3 2p^0 \delta^3(\vec{p}-\vec{q}) \delta^{rs}, \quad (10.12)$$

with other anti-commutators being zero. As a result, one finds

$$H = \int d\tilde{p} p^0 \sum_s (a_{\vec{p}}^{s\dagger} a_{\vec{p}}^s - b_{\vec{p}}^s b_{\vec{p}}^{s\dagger}) = \int d\tilde{p} p^0 \sum_s (a_{\vec{p}}^{s\dagger} a_{\vec{p}}^s + b_{\vec{p}}^{s\dagger} b_{\vec{p}}^s) + (\infty \mathbb{1}), \quad (10.13)$$

where the term proportional to unity is omitted, like in the scalar case, by declaring the Hamiltonian to be normal ordered.

Now let us have a look at the calculational details. Recall first the Hamiltonian

$$H = \int d^3x \bar{\psi} \left(-i\vec{\gamma}\vec{\nabla} + m \right) \psi \quad (10.14)$$

and the mode expansion of the field

$$\psi(\vec{x}) = \int d\vec{p} (a_{\vec{p}}^s u_s(p) e^{+i\vec{p}\vec{x}} + b_{\vec{p}}^{s\dagger} v_s(p) e^{-i\vec{p}\vec{x}}) \quad (10.15)$$

$$\bar{\psi}(\vec{x}) = \int d\vec{p}' (a_{\vec{p}'}^{s'\dagger} \bar{u}_{s'}(p') e^{-i\vec{p}'\vec{x}} + b_{\vec{p}'}^{s'} \bar{v}_{s'}(p') e^{+i\vec{p}'\vec{x}}). \quad (10.16)$$

We will use the standard relation

$$\int d^3x e^{i\vec{p}\vec{x} \pm i\vec{p}'\vec{x}} = (2\pi)^3 \delta^3(\vec{p} \pm \vec{p}') \quad (10.17)$$

in what follows.

In working out the Hamiltonian we find four terms, with $a^\dagger a$, $a^\dagger b^\dagger$, ba , and bb^\dagger respectively. Let us treat them one by one:

1.) The $a^\dagger a$ term reads

$$H_{a^\dagger a} = \int \frac{d\vec{p}}{2p^0} a_{\vec{p}}^{s'\dagger} a_{\vec{p}}^s \bar{u}_{s'}(p) (\vec{\gamma}\vec{p} + m) u_s(p). \quad (10.18)$$

We use

$$0 = (\not{p} - m)u(p) = (\gamma^0 p^0 - \vec{\gamma}\vec{p} - m)u(p), \quad (10.19)$$

from which it follows that

$$(\vec{\gamma}\vec{p} + m)u(p) = \gamma^0 p^0 u(p). \quad (10.20)$$

Thus, we find

$$H_{a^\dagger a} = \int \frac{d\vec{p}}{2} a_{\vec{p}}^{s'\dagger} a_{\vec{p}}^s \bar{u}_{s'}(p) \gamma^0 u_s(p). \quad (10.21)$$

A useful relation in this context is the following:

$$\bar{u}_r(p) \gamma^0 u_s(p) = \bar{v}_r(p) \gamma^0 v_s(p) = 2p^0 \delta_{rs}. \quad (10.22)$$

To prove it, we recall that

$$(\not{p} - m)u(p) = 0 \quad (10.23)$$

and hence

$$0 = u^\dagger(p)(\not{p}^\dagger - m) = u^\dagger(p)(\not{p}^\dagger - m)\gamma^0 = \bar{u}(p)(\not{p} - m). \quad (10.24)$$

With this, we may write

$$\begin{aligned} \bar{u}_r(p) \gamma^0 u_s(p) &= \frac{1}{2m} \bar{u}_r(p) \{m, \gamma^0\} u_s(p) \\ &= \frac{1}{2m} \bar{u}_r(p) \{\not{p} - m + m, \gamma^0\} u_s(p) \\ &= \frac{1}{2m} \bar{u}_r(p) \{\not{p}, \gamma^0\} u_s(p) \\ &= \frac{p^0}{m} \bar{u}_r(p) u_s(p) = \frac{p^0}{m} 2m \delta_{rs} = 2p^0 \delta_{rs}. \end{aligned} \quad (10.25)$$

An analogous calculation can be performed with v and \bar{v} . One then obtains

$$H_{a^\dagger a} = \int d\vec{p} p^0 a_{\vec{p}}^{s\dagger} a_{\vec{p}}^s, \quad (10.26)$$

concluding our discussion of the $a^\dagger a$ contribution.

2.) Now we turn to the $a^\dagger b^\dagger$ term. This term vanishes,

$$H_{a^\dagger b^\dagger} = 0, \quad (10.27)$$

which follows from manipulations analogous to those used in case 1.) together with the identities

$$u_s^\dagger(p^0, -\vec{p}) v_r(p^0, \vec{p}) = 0 \quad \text{and} \quad v_s^\dagger(p^0, -\vec{p}) u_r(p^0, \vec{p}) = 0. \quad (10.28)$$

The readers are invited to derive these identities by themselves, using tools that are already at their disposal (this is also treated in the problems).

3.) By an analogous reasoning, one can also show that

$$H_{ba} = 0. \quad (10.29)$$

4.) Similarly to case 1.), one finds

$$\begin{aligned} H_{bb^\dagger} &= \int \frac{d\vec{p}}{2p^0} b_{\vec{p}}^{s'} b_{\vec{p}}^{s\dagger} \bar{v}_{s'}(p) (-\vec{\gamma}\vec{p} + m) v_s(p) \\ &= \int \frac{d\vec{p}}{2p^0} b_{\vec{p}}^{s'} b_{\vec{p}}^{s\dagger} \bar{v}_{s'}(p) (-\gamma^0 p_0) v_s(p) \\ &= \int d\vec{p} p^0 \left(-b_{\vec{p}}^s b_{\vec{p}}^{s\dagger} \right). \end{aligned} \quad (10.30)$$

Thus, combining all the results above, we have

$$H = \int d\vec{p} p^0 \left(a_{\vec{p}}^{s\dagger} a_{\vec{p}}^s + b_{\vec{p}}^s b_{\vec{p}}^{s\dagger} \right) + \text{irrelevant constant}. \quad (10.31)$$

We define the Fock space just like in the bosonic case:

$$|0\rangle, a_{\vec{p}}^{s\dagger} |0\rangle, b_{\vec{p}}^{s\dagger} |0\rangle, a_{\vec{p}}^{s\dagger} a_{\vec{q}}^{r\dagger} |0\rangle, \dots \quad (10.32)$$

As before, the vacuum obeys

$$a_{\vec{p}}^s |0\rangle = b_{\vec{p}}^s |0\rangle = 0 \quad \forall \vec{p}, s. \quad (10.33)$$

The crucial difference to the bosonic case arises from the anti-commutation relations, which ensure that creation and annihilation operators square to zero. Thus,

$$\left(a_{\vec{p}}^{s\dagger} \right)^2 |0\rangle = 0, \quad (10.34)$$

i.e. multiple particle states with identical quantum numbers never occur. In standard quantum mechanical terminology, this means that our particles are fermions. Hence the name ‘spin-statistics theorem’, which was mentioned earlier. Actually, in the

plane-wave case in Minkowski space the index p is continuous, such that one might naively think that $p = p'$ never occurs anyway. This is of course not true if one thinks carefully about the right physical questions. For example one may take the non-relativistic limit and derive the familiar antisymmetry of the quantum-mechanical two-particle wave-function in real space. To appreciate the importance of the constraint above (i.e. of the Pauli exclusion principle) in a very simple-minded way, directly in quantum field theory, one should think of a finite volume situation. Then \vec{p} is discrete and the relation $\left(a_{\vec{p}}^{s\dagger}\right)^2 |0\rangle = 0$ obviously excludes a significant part of the naively available low-energy states in the Fock space.

10.4 Time ordering, Green's functions, Dirac propagators

Our fundamental goal is to go through the complete logic of perturbative quantum field theory, from LSZ to Wick's theorem to Feynman rules, once again, but this time with anti-commuting rather than commuting fields. This is straightforward and the readers are invited to check as much of this they want, consulting also the relevant textbooks. But the results will be extremely similar and the only difference is easy to understand:

Namely, whenever we write some expression involving products of fields, we simply attach a relative minus sign to every term involving an odd permutation of fermionic fields relative to the corresponding bosonic formula. For example, the very definition of the anti-commutator is of this type:

$$[\phi_1, \phi_2] = \phi_1\phi_2 - \phi_2\phi_1 \quad (10.35)$$

turns into

$$\{\psi_1, \psi_2\} = \psi_1\psi_2 - \text{sgn}(\sigma)\psi_{\sigma(1)}\psi_{\sigma(2)} = \psi_2\psi_1 + \psi_2\psi_1, \quad (10.36)$$

where $\{\sigma(1), \sigma(2)\} = \{2, 1\}$ is the unique non-trivial permutation of two elements. Here we suppress spinor indices. As another example, consider the derivation property of the commutator that was important in proving Wick's theorem:

$$[\phi_1, \phi_2\phi_3] = [\phi_1, \phi_2]\phi_3 + \phi_2[\phi_1, \phi_3]. \quad (10.37)$$

The fermionic analogue of this is

$$[\psi_1, \psi_2\psi_3] = \{\psi_1, \psi_2\}\psi_3 + \text{sgn}(\sigma)\psi_{\sigma(1)}\{\psi_{\sigma(2)}, \psi_{\sigma(3)}\}, \quad (10.38)$$

where $\{\sigma(1), \sigma(2), \sigma(3)\} = \{2, 1, 3\}$. The reader should check explicitly that the above relation is correct.

We note that the product of an even number of fermionic objects counts as bosonic while the product of a odd number of fermionic objects remains fermionic. Thus, for example, the derivation property applied to a product of three fermionic fields is formulated using an anti-commutator instead of a commutator: $\{\psi_1, \psi_2\psi_3\psi_4\}$. We will not spell this out in detail since the basic idea should by now be clear.

One of the basic objects of interest in our context is

$$\langle 0|T(\text{product of } \psi\text{'s and } \bar{\psi}\text{'s})|0\rangle. \quad (10.39)$$

To be able to use this object in analogy to the bosonic case, we need to change its definition as explained above. Thus, for fermionic fields we define:

$$T\psi_{a_1}(x_1)\dots\psi_{a_n}(x_n) \equiv \text{sgn}(\sigma)\psi_{a_{\sigma(1)}}(x_{\sigma(1)})\dots\psi_{a_{\sigma(n)}}(x_{\sigma(n)}) \quad (10.40)$$

where $\{\sigma(1)\dots\sigma(n)\}$ is a permutation σ of $\{1\dots n\}$ such that $x_{\sigma(1)}^0 \geq \dots \geq x_{\sigma(n)}^0$ and $\text{sgn}(\sigma) = \pm 1$ for even and odd σ , respectively. This generalizes in an obvious way to products of $\bar{\psi}$'s or combinations of $\psi, \bar{\psi}$.

With this definition of T , the LSZ-formula can be derived in complete analogy to the familiar bosonic case. An overall sign can arise if the order of the fields in the time-ordered correlation function does not match the order of creation/annihilation operators used to define the incoming/outgoing combinations of particles.

Moreover, the relation between time-ordered Green's functions for interacting and for free fields (with the crucial factor $\exp(iS_{int})$ under the T -symbol) still holds:

$$\langle 0|T\psi_1\dots\psi_n|0\rangle = \frac{{}_0\langle 0|T\psi_1^I\dots\psi_n^I\exp(iS_{int}^I)|0\rangle_0}{{}_0\langle 0|T\exp(iS_{int}^I)|0\rangle_0}. \quad (10.41)$$

The reader is invited to check that nothing changes in the derivation. One important aspect to note here is that S_{int} and H_{int} are always bosonic. Another relevant comment is that the order of operators should remain the same on both sides of all relations, even under the T -symbols. Otherwise, extra signs can arise.

Finally, in the last step towards the Feynman rules we need the fermionic version of the Wick theorem. The latter is affected by extra signs, as explained in the beginning of this section. Formally, the Wick theorem looks exactly as before:

$$T(\psi_1\dots\bar{\psi}_n) = :(\psi_1\dots\bar{\psi}_n + \text{all possible contractions}): \quad (10.42)$$

But in detail there are some crucial modifications. First, the meaning of normal-ordering has to be adjusted for fermions in a fairly obvious way: In a normal ordered product, all creators appear on the left of all annihilators, and an overall sign is included if the permutation of fermionic fields required to achieve this is odd.

Second, the definition of 'contraction' also includes a factor of (-1) for each exchange of neighbouring $\psi, \bar{\psi}$ -fields required to place contracted pairs next to each other. Going back to our derivation of Wick's theorem, one may easily convince oneself that these adjustments in fact necessary for the theorem to remain correct.

We give a simple example for how the factors (-1) (in the present case only a single such factor) appear:

$$:\overline{\psi_1\psi_2\psi_3}\bar{\psi}_4:= -\overline{\psi_1\psi_3}:\psi_2\bar{\psi}_4: \quad (10.43)$$

A contraction of two fields is defined in the same way as in the bosonic case:

$$\overline{\psi_a(x)\bar{\psi}^b(y)} \equiv \langle T\psi_a(x)\bar{\psi}^b(y) \rangle \equiv S_F(x-y)_a^b. \quad (10.44)$$

Contractions of $\psi\psi$ or $\bar{\psi}\bar{\psi}$ vanish or, if you wish, do not exist. This is analogous to the situation with a complex scalar, where the only non-vanishing contractions are those of ϕ with $\bar{\phi}$.

The Dirac propagator S_F (where the index ‘ F ’ denotes the Feynman- $i\epsilon$ -prescription) just introduced explicitly reads

$$S_F(x-y) = \int \frac{d^4p}{(2\pi)^4} \frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon} e^{-ip(x-y)}. \quad (10.45)$$

The derivation of this is completely analogous to the bosonic case. The only difference is that a factor $\not{p} + m$ in the numerator arises from the u ’s and v ’s which are part of the mode expansion of the spinor fields.

We do not go through this straightforward but somewhat tedious analysis. Instead, we provide an alternative argument which is much simpler. It does not look the moment like a proper derivation but rather like a way to guess the correct result. However, using the path or functional integral approach studied in Quantum Field Theory II, it will become clear that it is, in fact, completely trustworthy and rigorous.

We start by recalling that, in the scalar case,

$$-(\square_x + m^2)D(x-y) \equiv i\delta^4(x-y). \quad (10.46)$$

Any D satisfying this relation is a Green’s function. It is not unique since the Klein-Gordon operator has zero modes. Feynman’s $i\epsilon$ prescription fixes this ambiguity:

$$D_F(x-y) = \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2 + i\epsilon} e^{-ip(x-y)}. \quad (10.47)$$

Crucially, the denominator $p^2 - m^2$ is just the Fourier-space version of $-(\square_x + m^2)$. Depending on the pole-prescription used, we get the Feynman, retarded or advanced Green’s function.

Analogously, the Green’s function S for spinors is defined by

$$(i\not{\partial}_x - m)S(x-y) \equiv \mathbb{1} i\delta^4(x-y). \quad (10.48)$$

The solution with Feynman prescription can be read off immediately:

$$S_F(x-y) = \int \frac{d^4p}{(2\pi)^4} \frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon} e^{-ip(x-y)}. \quad (10.49)$$

One simply has to take the Dirac operator to Fourier space,

$$(i\not{\partial} - m) \rightarrow (\not{p} - m), \quad (10.50)$$

and to find the inverse:

$$(\not{p} - m) \frac{\not{p} + m}{p^2 - m^2} = \mathbb{1}. \quad (10.51)$$

Here we have used that $(\not{p} + m)(\not{p} - m) = (p^2 - m^2) \mathbb{1}$.

Equivalently, we may directly write down the inverse of the Dirac operator in Fourier space, $1/(\not{p} - m)$, and reformulate it in a way which makes it obvious how to introduce the $i\epsilon$ by analogy to the bosonic case:

$$\frac{1}{\not{p} - m} = \frac{\not{p} + m}{p^2 - m^2} \rightarrow \frac{\not{p} + m}{p^2 - m^2 + i\epsilon}. \quad (10.52)$$

10.5 U(1)-Symmetry of the Dirac lagrangian

Consider the free Dirac-Lagrangian

$$\mathcal{L} = \bar{\psi}(i\cancel{\partial} - m)\psi \quad (10.53)$$

and observe that it is invariant under the *global* symmetry transformation:

$$\psi \rightarrow e^{-i\epsilon}\psi \simeq \psi - i\epsilon\psi. \quad (10.54)$$

Moreover, recall Noether's theorem states that

$$j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\chi - F^\mu \quad (10.55)$$

is a conserved current if

- (1) if $\varphi \rightarrow \varphi + \epsilon\chi$ is the infinitesimal version of a continuous global symmetry transformation and
- (2) the lagrangian transforms as $\mathcal{L} \rightarrow \mathcal{L} + \epsilon\partial_\mu F^\mu$.

In the case at hand we obtain

$$j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}(-i\psi) = \bar{\psi}i\gamma^\mu(-i\psi) = \bar{\psi}\gamma^\mu\psi, \quad (10.56)$$

where we have used that $F^\mu = 0$ and $\chi = -i\psi$.

Jumping ahead, we note that this current will become the electromagnetic current after gauging the $U(1)$ or, in other words, making the $U(1)$ -symmetry 'local'.

Of course, with our current comes a conserved charge Q . It can be evaluated straightforwardly using the relations just obtained and the mode decomposition of our field:

$$Q = \int d^3x j^0 = \int d^3x \psi^\dagger\psi = \int d\tilde{p} \sum_s \left(a_{\tilde{p}}^{s\dagger} a_{\tilde{p}}^s - b_{\tilde{p}}^{s\dagger} b_{\tilde{p}}^s \right). \quad (10.57)$$

10.6 Yukawa theory

The arguably simplest interacting theory with fermions is the so-called Yukawa theory. In its minimalist version, it is described by the lagrangian

$$\mathcal{L} = \underbrace{\bar{\psi}(i\cancel{\partial} - m)\psi}_{\text{free fermion}} + \underbrace{\frac{1}{2}(\partial\varphi)^2 - \frac{\mu^2}{2}\varphi^2}_{\text{free scalar}} + \underbrace{\lambda\varphi\bar{\psi}\psi}_{\text{Yukawa interaction}}. \quad (10.58)$$

Interactions of this type play a fundamental role for the fermion-mass generation in the Standard Model. There, $m = 0$ and the real scalar φ is replaced by the Higgs boson (a 2-component complex scalar). The latter has a non-zero value in the vacuum (a so-called vacuum expectation value or VEV) providing a mass for the fermions which is proportional to the coupling constant λ . This so-called Yukawa coupling is different for each fermion and responsible for their different masses.

The Yukawa theory also plays a phenomenological role in nuclear physics. There, the relevant fermions are proton and neutron and the exchanged scalars are the pion fields. Both these fermions and the scalars are built from more fundamental objects, the quarks and gluons of Quantum Chromodynamics or QCD. Hence, in this case the Yukawa theory is not **fundamental** but only provides an **effective** description. Of course, also in the Standard Model case it is possible that, probing quarks, leptons and Higgs field at higher energies we will discover that these and hence their Yukawa-type interactions are not fundamental either.

Some more details of the Yukawa interactions will be developed in the problems. We will now turn to the more complicated and structurally more interesting **gauge interactions** of fermions.

11 Quantum Electrodynamics

11.1 Lagrangian

The logic underlying the QED Lagrangian is the same as in ‘scalar QED’ discussed earlier: We promote the global U(1)-symmetry of

$$\mathcal{L} = \bar{\psi}(i\not{\partial} - m)\psi \quad (11.1)$$

to a local symmetry:

$$\psi \rightarrow e^{-i\alpha(x)}\psi. \quad (11.2)$$

In order to maintain the invariance of our action under this larger ‘gauge’ symmetry, we also need to promote the partial derivative in the free-fermion lagrangian to a covariant derivative:

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + iA_\mu. \quad (11.3)$$

Finally, a kinetic term for the gauge potential A_μ has to be introduced.

Thus, the QED Lagrangian reads

$$\mathcal{L}_{\text{QED}} = -\frac{1}{4e^2} F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\not{D} - m)\psi, \quad (11.4)$$

where $\not{D} \equiv \gamma^\mu D_\mu$. Alternatively, by redefining the gauge field A we may write

$$\mathcal{L}_{\text{QED}} = -\frac{1}{4} F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\not{D} - m)\psi \quad (11.5)$$

with $D_\mu = \partial_\mu + ieA_\mu$. Let us check the claimed gauge-invariance explicitly: We have

$$\begin{aligned} D_\mu\psi &\rightarrow D'_\mu\psi' = (\partial_\mu + ieA'_\mu)e^{-ie\alpha(x)}\psi = e^{-ie\alpha(x)}(\partial_\mu - ie\partial_\mu\alpha + ieA'_\mu)\psi \\ &= e^{-ie\alpha(x)}D_\mu\psi, \end{aligned} \quad (11.6)$$

where we used $A'_\mu = A_\mu + \partial_\mu\alpha(x)$ in the last step. The exponential factor drops out when the derivative of the spinor ψ is combined with its counterpart $\bar{\psi}$ in the Lagrangian, which is hence indeed invariant.

11.2 Deriving the Feynman rules

We split the Lagrangian according to

$$\mathcal{L}_{\text{QED}} = \mathcal{L}_{\text{free}} + \mathcal{L}_{\text{int}}, \quad (11.7)$$

with the free Lagrangian

$$\mathcal{L}_{\text{free}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\psi}(i\not{\partial} - m)\psi \quad (11.8)$$

and the interaction Lagrangian

$$\mathcal{L}_{\text{int}} = -e\bar{\psi}\not{A}\psi. \quad (11.9)$$

The interaction Lagrangian arises exclusively from the so-called **minimal coupling** that is needed to ensure gauge-invariance.

As a side remark, we note that a **non-minimal coupling** would, for example, be

$$\mathcal{L}_{\text{int}} = \frac{1}{\Lambda} F_{\mu\nu} \bar{\psi} \gamma^\mu \gamma^\nu \psi. \quad (11.10)$$

It is suppressed by a mass-dimension parameter, which is typically related to the cutoff scale Λ (we have identified the two). This coupling is less important if Λ is large. Also, this coupling is forbidden if we require our theory to be renormalisable - a term characterising theories where infinities can be removed in controlled way (to be discussed further down).

We note that our bosonic fields and (obviously) the partial derivative have **mass dimension** one. This is written as

$$[\varphi] = [A_\mu] = [\partial_\mu] = E^1, \quad (11.11)$$

where $[\cdot]$ denotes the power of *mass* or *energy* which has the same unit as the object inside the brackets. Recall that we have set $\hbar = c = 1$.

From this, it follows that the fermion field has mass dimension $3/2$, i.e.

$$[\psi] = E^{3/2}. \quad (11.12)$$

To see this, note that

$$[S] = [\hbar] = E^0; \quad [d^4x] = E^{-4} \quad \text{and hence} \quad [\mathcal{L}] = E^4. \quad (11.13)$$

Combining this with (11.8) & (11.11) we obtain (11.12). Also, from (11.10) it follows that:

$$[\Lambda] = E, \quad (11.14)$$

a fact which we already stated without prove above. Typically, Λ is the energy scale at which some **new physics** (beyond the QED lagrangian) appears. This new physics might be responsible for or 'generate' the non-minimal coupling we discussed. One also says that this non-minimal coupling corresponds to a **higher-dimension operator**, i.e. an expression the mass dimension of which is larger than four. In our

case $[F_{\mu\nu}\bar{\psi}\gamma^\mu\gamma^\nu\psi] = E^5$. We now return to pure QED without higher-dimension operators.

For the fermion, the free lagrangian $\mathcal{L}_{\text{free}}$ implies the following Feynman rule:

$$a \begin{array}{c} \longrightarrow \\ \text{p} \end{array} b = \left(\frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon} \right)_b^a. \quad (11.15)$$

As before, this is just the Fourier-space expression for $\langle T\psi_b(x)\bar{\psi}^a(y) \rangle$. As a side remark, we note that it is equivalent to add the $i\epsilon$ directly to the matrix in the denominator:

$$\begin{aligned} \frac{i}{\not{p} - m + i\epsilon} &= \frac{i(\not{p} + m - i\epsilon)}{(\not{p} + m - i\epsilon)(\not{p} - m + i\epsilon)} = \frac{i(\not{p} + m - i\epsilon)}{p^2 - (m - i\epsilon)^2} \\ &= \frac{i(\not{p} + m - i\epsilon)}{p^2 - m^2 + 2mi\epsilon + \epsilon^2} \hat{=} \frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon'}. \end{aligned} \quad (11.16)$$

In the last step we neglected the ϵ in the numerator as we are only interested in the poles. Furthermore, in the small ϵ -limit we can neglect the ϵ^2 -term. By redefining $\epsilon' = 2m\epsilon \rightarrow \epsilon$ we obtain the already known expression for the propagator in Fourier space.

In our simplest gauge choice (Feynman gauge) the Feynman rule for the gauge propagation is

$$\mu \begin{array}{c} \text{~~~~~} \\ \text{p} \rightarrow \end{array} \nu = \frac{-i\eta^{\mu\nu}}{p^2 + i\epsilon}. \quad (11.17)$$

Finally, the interaction lagrangian \mathcal{L}_{int} implies the Feynman rule

$$\begin{array}{c} b \\ \nearrow \\ \text{~~~~~} \\ \searrow \\ a \end{array} \mu = -ie(\gamma_\mu)_b^a. \quad (11.18)$$

Up to a possible sign this should be clear: The vertex is just the coefficient of the 3-field term in \mathcal{L} . To derive this more carefully, let us consider an imagined process $e^+ + \gamma \rightarrow e^+$ (with momenta $p + k = p'$). For the corresponding matrix element we have

$$\langle 0 | a_{\vec{p}'}^{s'} \left(i \int d^4x \mathcal{L}_{\text{int}} \right) a_{\vec{p}}^{s\dagger} a_{\vec{k}}^{\mu\dagger} | 0 \rangle \left(-\epsilon_\mu(k) \right) = (2\pi)^4 \delta^4(\dots) i\mathcal{M}_{\text{fi}}, \quad (11.19)$$

where the annihilation/creation operators account for:

$a_{\vec{p}'}^{s'}$: outgoing positron with spin s' .

$a_{\vec{p}}^{s\dagger}$: incoming positron with spin s .

$a_{\vec{k}}^{\mu\dagger}$: incoming photon with polarisation $\epsilon_\mu(k)$.

Thus, we now know how a vertex connects to incoming and outgoing states. Next, we will see how it connects to internal lines, i.e. to propagators. For this purpose, consider the process

$$e^+\gamma \rightarrow e^+\gamma : \quad \begin{array}{c} \text{---} \\ \diagup \quad \diagdown \\ \text{---} \end{array} \quad , \quad (11.26)$$

with time flowing from left to right in the diagram. The corresponding amplitude follows from the second-order term in the perturbative series in the interaction lagrangian:

$$\langle 0 | a^{s'} a^{\mu'} T \left[\left(\int_x \bar{\psi}(-ie\gamma_\nu A^\nu) \psi \right) \left(\int_y \bar{\psi}(-ie\gamma_\rho A^\rho) \psi \right) \right] a^{s\dagger} a^{\mu\dagger} | 0 \rangle . \quad (11.27)$$

Here contractions involving a 's are meant as an informal symbol for producing a non-zero number by commutation relations, similar to the actual contractions. A detailed evaluation of this expression gives

$$\begin{array}{c} k \quad k' \\ \diagdown \quad \diagup \\ \text{---} \\ \diagup \quad \diagdown \\ p \quad p' \\ \text{---} \\ \text{Read from left to right} \end{array} = \underbrace{\varepsilon_{\mu'}^*(k') \bar{u}_{s'}(p') (ie\gamma^{\mu'}) \frac{i}{\not{q} - m + i\varepsilon} (ie\gamma^\mu) u_s(p) \varepsilon_\mu(k)}_{\text{Read from right to left}} , \quad (11.28)$$

where the contraction of matrix indices from right to left corresponds to moving along the fermion line, in the direction of the arrow. As an important technical conclusion we note: Our definition of the propagator (where the arrow corresponds to going from $\bar{\psi}$ to ψ) is consistent with the way the arrow was introduced in (11.24) for external particles.

Finally, let us consider external antiparticles, in this case electrons. We study the (unrealistic) process $e^-\gamma \rightarrow e^-$ with momenta $p+k = p'$. For the matrix element we have

$$\langle 0 | b_{s'} \left(\int \bar{\psi}(-ieA) \psi \right) b_s^\dagger a_\mu^\dagger | 0 \rangle , \quad (11.29)$$

and thus

$$\begin{array}{c} p \\ \diagdown \\ \text{---} \\ \diagup \\ k \\ \text{---} \\ \text{read from left to right} \end{array} \rightarrow p' = \underbrace{\varepsilon_\mu(k) \bar{v}_s(p) (-ie\gamma^\mu) v_{s'}(p')}_{\text{time flows also from left to right}} . \quad (11.30)$$

Summarising we can say: In the string of matrices corresponding to a fermion line, time flows from right to left for particles and from left to right for antiparticles.

11.3 Summary of QED Feynman rules and simple examples

Let us summarise the QED Feynman rules:

$$\text{---}\blacktriangleright\text{---} = \frac{i}{\not{k} - m + i\epsilon} \quad (11.31)$$

$$\text{~~~~~} = \frac{-i\eta^{\mu\nu}}{k^2 + i\epsilon} \quad (11.32)$$

$$\begin{array}{c} b \\ \swarrow \\ \text{~~~~~} \\ \nwarrow \\ a \end{array} = -ie(\gamma^\mu)_b^a . \quad (11.33)$$

We include the rules for external fermions:

$$p \rightarrow \text{---}\blacktriangleright\text{---}\text{---}\text{---} = (\dots)u(p) \quad \text{incoming particle} \quad (11.34)$$

$$p \rightarrow \text{---}\blacktriangleleft\text{---}\text{---}\text{---} = \bar{v}(p)(\dots) \quad \text{incoming antiparticle}$$

$$\text{---}\text{---}\blacktriangleright\text{---} \rightarrow p = \bar{u}(p)(\dots) \quad \text{outgoing particle} \quad (11.35)$$

$$\text{---}\text{---}\blacktriangleleft\text{---} \rightarrow p = (\dots)v(p) \quad \text{outgoing antiparticle} .$$

Of course, as in the bosonic case, Z -factors have to be added. We have suppressed them for brevity. The Z -factors are the same for particle and antiparticle.

There are some **additional sign rules** characteristic of fermion diagrams: A diagram receives a relative minus sign for

1. Every closed fermion line (loop).
2. The exchange of two external fermion lines (relative to another diagram).

We will not derive them in generality but merely illustrate how they arise in two examples:

Consider the following two diagrams for $e^+e^- \rightarrow e^+e^-$:

$\langle 2'1'(\bar{\psi}A\psi)(\bar{\psi}A\psi)(\bar{\psi}A\psi)(\bar{\psi}A\psi)12 \rangle$

$\langle 2'1'(\bar{\psi}A\psi)(\bar{\psi}A\psi)(\bar{\psi}A\psi)(\bar{\psi}A\psi)12 \rangle$

(11.36)

They are of order e^4 for the amplitude. Since a non-zero amplitude for this process arises already at the order e^2 , and no e^3 contribution exists, one calls such a correction an ‘NLO contribution’, with NLO standing for ‘next-to-leading order’.

Each of these two diagrams acquired two minus signs (and hence no minus in total) from ‘intersecting fermion contractions’. But, crucially, in addition the right-hand side has a contraction $\overline{\psi\psi} = -\overline{\psi\psi} = -S_F$. We claim (and the reader should check using other examples and thinking about the structure contractions) that such an extra minus *always* occurs if there is a closed fermion loop.

The minus sign associated with the exchange of external lines comes simply from

$$\langle 0| a_{\vec{k}'}^r a_{\vec{p}'}^s = - \langle 0| a_{\vec{p}'}^s a_{\vec{k}'}^r . \quad (11.37)$$

It is easy to see that this relative minus sign appears for example in

(11.38)

11.4 Elementary processes

In this section, we will mention a few famous QED scattering processes and then pick a particularly simple one to work out in detail.

First, there is **Compton scattering**: $e^- \gamma \rightarrow e^- \gamma$. The relevant LO (leading-order) diagrams are

$$\begin{array}{c} \text{Diagram 1} \end{array} + \begin{array}{c} \text{Diagram 2} \end{array}, \quad (11.39)$$

and the resulting cross section is described by the **Klein-Nishina formula**.

Then there is **Møller scattering**: $e^- e^- \rightarrow e^- e^-$ (or $e^+ e^+ \rightarrow e^+ e^+$), described by the diagrams

$$\begin{array}{c} \text{Diagram 1} \end{array} + \begin{array}{c} \text{Diagram 2} \end{array}. \quad (11.40)$$

A closely related process is that in which the two scattering particles are distinguished, like e.g. $e^- \mu^- \rightarrow e^- \mu^-$. Then clearly only the first diagram contributes.

Of this last process, one may now consider two particularly interesting limiting cases: The first is the case of a non-relativistic target and a highly relativistic projectile. At high but not too high energy, this is automatically realised if, for example, an electron scatters off a muon or off a nucleus. This is **Coulomb scattering**, where the cross section is characterised by the **Mott formula**. The name ‘Coulomb scattering’ implies scattering off a static Coulomb field.

The second case is the case of a non-relativistic projectile and a non-relativistic target. This is called **Rutherford scattering**, with the famous **Rutherford formula** describing the cross section.

Further interesting processes are **pair annihilation to photons**, $e^+ e^- \rightarrow \gamma \gamma$,

$$\begin{array}{c} \text{Diagram 1} \end{array} + \begin{array}{c} \text{Diagram 2} \end{array}, \quad (11.41)$$

Bhabha scattering, $e^+ e^- \rightarrow e^+ e^-$,

$$\begin{array}{c} \text{Diagram 1} \end{array} + \begin{array}{c} \text{Diagram 2} \end{array}, \quad (11.42)$$

and **light-by-light scattering**, $\gamma\gamma \rightarrow \gamma\gamma$,

An interesting special feature of the last process is that it does not occur as a tree-level process, only at loop-order.

As an example to work out in detail we pick the scattering process $e^+e^- \rightarrow \mu^+\mu^-$. This is similar to Bhabha scattering but even simpler since only the following single diagram is involved:

We have the kinematic relations

$$q = p + p' = k + k' \quad \text{and} \quad q^2 \equiv s \quad (11.45)$$

and, using our Feynman rules derived earlier, the invariant matrix element reads

$$i\mathcal{M} = \bar{u}_s(k)ie\gamma_\mu v_{s'}(k') \frac{-i\eta^{\mu\nu}}{q^2 + i\varepsilon} \bar{v}_{r'}(p')ie\gamma_\nu u_r(p). \quad (11.46)$$

As in section 5.3, the differential cross section in the limit $\sqrt{s} \gg m_e, m_\mu$ can be given as

$$d\sigma = \frac{1}{2s} |\mathcal{M}|^2 dX^{(2)} = \frac{1}{64\pi^2 s} |\mathcal{M}|^2 d\Omega. \quad (11.47)$$

Let us now assume that we have unpolarized incoming beams and that the spin of the outgoing particles is not measured. Then we need to average over incoming and sum over outgoing spins. This is realised by replacing the invariant matrix element squared in (11.47) according to

$$\begin{aligned} |\mathcal{M}|^2 &\rightarrow \underbrace{\frac{1}{2} \sum_r}_{\text{average}} \underbrace{\frac{1}{2} \sum_{r'} \sum_s \sum_{s'}}_{\text{sum}} |\mathcal{M}(r, r', s, s')|^2 \\ &= \frac{e^4}{4s^2} \sum_{s, s'} (\bar{u}_s(k) \gamma_\mu v_{s'}(k')) \overline{(\bar{u}_s(k) \gamma_\nu v_{s'}(k'))} \cdot \sum_{r, r'} (\bar{v}_{r'}(p') \gamma^\mu u_r(p)) \overline{(\bar{v}_{r'}(p') \gamma^\nu u_r(p))} \\ &\equiv \frac{e^4}{4s^2} A_{\mu\nu} B^{\mu\nu}. \end{aligned} \quad (11.48)$$

Let us evaluate the first of the two tensor factors into which we have split the result in the last line:

$$\begin{aligned}
A_{\mu\nu} &= \sum_{s,s'} \text{tr} [\bar{u}_s(k) \gamma_\mu v_{s'}(k') \bar{v}_{s'}(k') \gamma_\nu u_s(k)] \\
&= \sum_{s,s'} \text{tr} [u_s(k) \bar{u}_s(k) \gamma_\mu v_{s'}(k') \bar{v}_{s'}(k') \gamma_\nu] \\
&= \text{tr} [(\not{k} + m_\mu) \gamma_\mu (\not{k}' - m_\mu) \gamma_\nu] = \text{tr} [\not{k} \gamma_\mu \not{k}' \gamma_\nu] - m_\mu^2 \text{tr} [\gamma_\mu \gamma_\nu] \\
&= 4(k_\mu k'_\nu + k'_\mu k_\nu - (k \cdot k') \eta_{\mu\nu} - m_\mu^2 \eta_{\mu\nu}) .
\end{aligned} \tag{11.49}$$

Here, in the first line of (11.49), we interpreted the scalar expression as a 1×1 matrix, so the matrix is identical to its trace. To get to the second line we made use of the cyclicity of the trace. Then we applied the completeness relations (9.85) as well as the following trace identity of gamma matrices:

$$\text{tr} [\gamma_\mu \gamma_\nu \gamma_\rho \gamma_\sigma] = 4 (\eta_{\mu\nu} \eta_{\rho\sigma} + \eta_{\mu\sigma} \eta_{\nu\rho} + \eta_{\mu\rho} \eta_{\nu\sigma}) . \tag{11.50}$$

A completely analogous calculation can be performed for $B_{\mu\nu}$.

From now on we neglect the m^2 -terms in $A_{\mu\nu}$ and $B^{\mu\nu}$, since $m_\mu, m_e \ll \sqrt{s}$. We find:

$$\begin{aligned}
\frac{1}{4} \left(\sum \right)^4 |\mathcal{M}|^2 &= \frac{e^4}{4s^2} 16 (k_\mu k'_\nu + k'_\mu k_\nu - \eta_{\mu\nu} (k \cdot k')) (p^\mu p'^\nu + p'^\mu p^\nu - \eta^{\mu\nu} (p \cdot p')) \\
&= \frac{8e^4}{s^2} ((k \cdot p)(k' \cdot p') + (k \cdot p')(k' \cdot p)) = 2e^4 \frac{t^2 + u^2}{s^2} .
\end{aligned} \tag{11.51}$$

Here, in the last expression we have introduced the so-called the **Mandelstam variables** s, t, u . These are defined by

$$s = (p + p')^2, \quad t = (p - k)^2, \quad u = (p - k')^2 \tag{11.52}$$

and they satisfy the relation $s + t + u = \sum_{i=1}^4 m_i^2$. To get some intuition, consider the momenta as defined in Fig. 13 and note that \sqrt{s} is the energy flowing in a left-to-right interpretation of this figure as a scattering process. By contrast, \sqrt{t} corresponds to the energy flowing in a top-to-bottom interpretation. Finally, \sqrt{u} is the energy flowing if we re-interpret this figure with p and $-k'$ as incoming momenta.

In the massless case, i.e. with

$$p^2 = p'^2 = k^2 = k'^2 = 0, \tag{11.53}$$

we have the following relations:

$$s = 2pp' = 2kk', \quad t = -2kp = -2k'p', \quad u = -2pk' = -2kp'. \tag{11.54}$$

Let us go to the center of mass system (cms) and express these quantities using the scattering angle, cf. Fig. 14:

$$t = -2kp = -2(k_0 p_0 - \vec{k} \vec{p}) = -2k_0 p_0 (1 - \cos \Theta) \tag{11.55}$$

$$= -2 \left(\frac{\sqrt{s}}{2} \right)^2 (1 - \cos \Theta) = -\frac{s}{2} (1 - \cos \Theta). \tag{11.56}$$

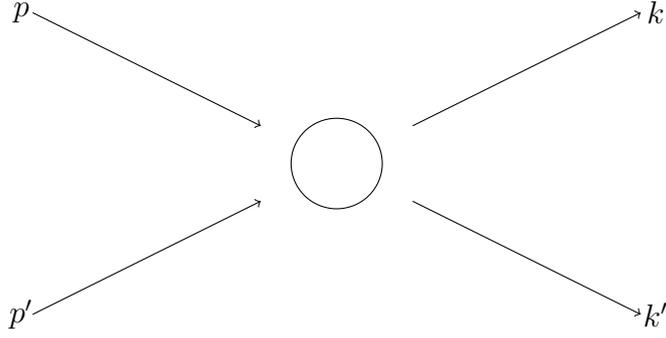


Figure 13: Momenta in a scattering process drawn to illustrate the meaning of s , t and u . See explanations in the text.

Next, using the sum rule for s , t and u , we have

$$u = -s - t = -\frac{s}{2}(1 + \cos \Theta). \quad (11.57)$$

Inserting this in (11.51) we find

$$\frac{1}{4} \left(\sum \right)^4 |\mathcal{M}|^2 = e^4(1 + \cos^2 \Theta). \quad (11.58)$$

Hence the cross section reads

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2 s} \frac{1}{4} \left(\sum \right)^4 |\mathcal{M}|^2 = \frac{\alpha^2}{4s}(1 + \cos^2 \Theta), \quad (11.59)$$

where $\alpha \equiv e^2/4\pi$.

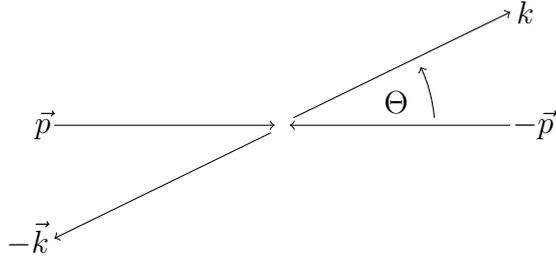


Figure 14: Definition of the scattering angle in the cms.

We notice that the angular dependence encodes an interesting phenomenon, illustrated in Fig. 15: Small scattering angles are preferred relative to large ones. We will next try to derive this preference in an intuitive way. To do so, let us recall that $\sum |\mathcal{M}|^2$ may be written

$$\sum_{\text{spins}} (\bar{v}\gamma^\mu u)(\bar{v}\gamma^\nu u) \eta^{\mu\mu'} \eta^{\nu\nu'} \sum_{\text{spins}} (\bar{u}\gamma^{\mu'} v)(\bar{u}\gamma^{\nu'} v) = B_{\mu\nu} \eta^{\mu\mu'} \eta^{\nu\nu'} A_{\mu'\nu'}. \quad (11.60)$$

In this expression, we may perform the substitution

$$\eta^{\mu\mu'} \rightarrow \eta^{\mu\mu'} - \frac{q^\mu q^{\mu'}}{q^2}, \quad (11.61)$$

and similarly for $\eta^{\nu\nu'}$. The reason is that $q = p + p' = k + k'$ and that, due to relations like $\not{p}u(p) = 0$, $\bar{u}(k)\not{k} = 0$ etc., the term $\sim q^\mu q^{\mu'}$ drops out.

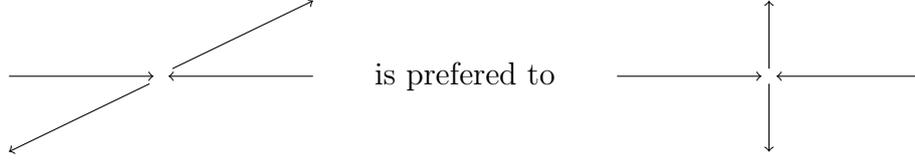


Figure 15: Illustration of intuitive meaning of the angular dependence.

As a result, in the center-of-mass system only the spatial part of $\eta^{\mu\mu'}$ and $\eta^{\nu\nu'}$ remains non-zero. In other words, we may replace

$$\eta^{\mu\mu'} \rightarrow \delta^{ii'} \quad \text{and} \quad \eta^{\nu\nu'} \rightarrow \delta^{jj'}. \quad (11.62)$$

This implies that also of the matrices A, B only the spatial parts are relevant. We may replace them according to

$$B_{\mu\nu} \rightarrow B_{ij} \quad \text{and} \quad A_{\mu'\nu'} \rightarrow A_{i'j'}. \quad (11.63)$$

Let us finally note that $\bar{v}(p')\gamma_\mu u(p)$ can be viewed as the polarization of the spin-1 state created from the two incoming particles. Hence, we may think of

$$\sum_{r'r} (\bar{v}_{r'}(p')\gamma_\mu u_r(p)) \overline{(\bar{v}_{r'}(p')\gamma_\nu u_r(p))} = B_{\mu\nu} \rightarrow B_{ij} \sim \rho_{ij}^{\text{ini}} \quad (11.64)$$

as of the quantum mechanical density matrix characterising the polarisation of this initial state. The concept of a density matrix comes in since, due to the unpolarised beam, our information is incomplete. Analogously, we may think of the spatial part of $A_{\mu\nu}$ as of a final-state density matrix:

$$A_{\mu\nu} \rightarrow A_{ij} \sim \rho_{ij}^{\text{fin}}. \quad (11.65)$$

In this language, we have

$$\sum |\mathcal{M}|^2 \sim \sum_{ij} \rho_{ij}^{\text{ini}} \rho_{ij}^{\text{fin}}. \quad (11.66)$$

We also note that the indices i, j correspond to the three physical polarizations of the intermediate, massive photon.

As we know from our explicit calculation, we have $B^{\mu\nu} \sim p^\mu p'^\nu + p'^\mu p^\nu - \eta^{\mu\nu}(pp')$. Rotating the cms in such a way that $p = (0, 0, 1)^T$, one thus finds

$$\rho_{\text{ini}}^{ij} \sim \delta^{ij} - \hat{p}^i \hat{p}^j \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}^{ij}, \quad (11.67)$$

where $\hat{p} \equiv \vec{p}/|\vec{p}|$. For the final state we have an analogous but in general rotated matrix:

$$\rho_{\text{fin}}^{ij} \sim \delta^{ij} - \hat{k}^i \hat{k}^j. \quad (11.68)$$

The angular dependence arising in this language is of course the same that we found before:

$$\sum |\mathcal{M}|^2 \sim \text{tr}(\rho_{\text{ini}}\rho_{\text{fin}}) \sim 3 - 1 - 1 + (\hat{k}\hat{p})^2 = 1 + \cos^2 \Theta. \quad (11.69)$$

But we can now understand more clearly where it comes from: As we see from the form of the density matrices, the two incoming spin-1/2 particles always produce a photon with spin ± 1 , never one with spin 0 along the 3-axis. Indeed, recall that the relevant polarisation vectors are proportional to

$$\begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (11.70)$$

The last one does clearly not contribute to our initial-state density matrix above. Analogous statements apply to the final-state density matrix. Due to this missing third photon polarisation, the trace of the product of the two matrices is sensitive to rotations. If each were a unit matrix, representing all three photons polarisations, no angular dependence would arise.

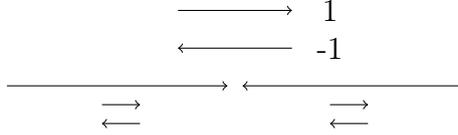


Figure 16: Spinor polarisations adding up to produce the photon polarisation.

We may also express what was said above by arguing about how the spinor polarisations combine to form the polarisation of the photon, cf. Fig. 16. The two incoming spinors make a polarisation $+1$ or -1 photon, along the beam axis. The outgoing spinors analogously make a polarisation $+1$ or -1 photon along the decay axis. Hence the correlation between beam and decay axis.

But **why** is the mathematically also possible combination

$$(+1/2)_{\text{fermion}} + (-1/2)_{\text{fermion}} = 0_{\text{photon}} \quad (11.71)$$

not realised? Naively, one would think that this is also a perfectly good process for producing a massive intermediate photon.

The reason is that we are in the high-energy limit, where the fermion mass can be neglected. Without a mass term, we are in fact dealing with two independent and uncoupled fields. Indeed, as one can easily check, the Dirac lagrangian may be rewritten according to

$$\bar{\psi}(i\mathcal{D} - m)\psi = \bar{\psi}_L i\mathcal{D}\psi_L + \bar{\psi}_R i\mathcal{D}\psi_R - m\bar{\psi}_L\psi_R - m\bar{\psi}_R\psi_L. \quad (11.72)$$

If one drops the mass term, this means that (as long as one does not use the photon propagator), there are two independent theories: one for the l.h. and one for the r.h. field. The resulting possibilities for an annihilation process are illustrated in Fig. 17. We see that no spin-0 photon (along the 3-axis) can be produced. The underlying reason is that **no helicity flip** is possible in the absence of a mass-term. More precisely, it is suppressed in the high-energy limit, when the mass is negligible.

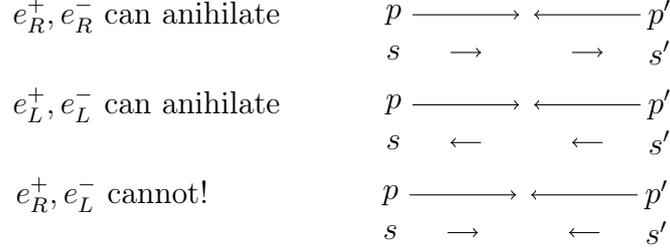


Figure 17: Possible annihilation processes in the absence of a mass term.

12 Renormalization

12.1 Concept

Let us denote by $\{Q_j, j = 1, 2, \dots\}$ the set of quantities we would like to calculate in a given QFT. Typically, this might be cross-sections, decay rates, various Greens-functions in real or Fourier space etc. Let us call them **observables**, although to some of them, like Green's functions, our access is somewhat indirect. Performing the necessary calculations to obtain such observables in perturbation theory one encounters, at higher orders, divergent loop integrals. A simple example is the amplitude for 2-to-2-scattering in $\lambda\varphi^4$ -theory at NLO:

$$\begin{array}{c}
 \text{Diagram 1} + \text{Diagram 2} + \dots \\
 \downarrow \\
 \int d^4k \frac{1}{k^2 - m^2 + i\epsilon} \cdot \frac{1}{(k+q)^2 - m^2 + i\epsilon}
 \end{array} \tag{12.1}$$

Let us, for the moment, **regularise** by analytically continuing ('Wick rotating') to Euclidean space, $k^2 \rightarrow k_E^2 = k_0^2 + k_1^2 + k_2^2 + k_3^2$, and introducing a **cutoff** in our integration: $|k_E| < \Lambda$. As a result, we have $Q_j = Q_j(\Lambda)$ and naively the limit $\Lambda \rightarrow \infty$ cannot be taken. **Renormalization** is the method to properly taking this limit nevertheless.

We will explain this method using our main example, QED. We write the lagrangian as

$$\mathcal{L} = -\frac{1}{4}F_{0,\mu\nu}F_0^{\mu\nu} + \bar{\psi}_0(i(\not{\partial} + ie_0\not{A}_0) - m_0)\psi_0 \tag{12.2}$$

where $F_0^{\mu\nu} = \partial^\mu A_0^\nu - \partial^\nu A_0^\mu$. The index 0 denotes so called **bare** fields and couplings. Next, we rewrite \mathcal{L} in terms of **renormalised** quantities:

$$A_0^\mu = Z_A^{1/2} A^\mu ; \quad \psi_0 = Z_\psi^{1/2} \psi ; \quad e_0 = Z_e e ; \quad m_0 = Z_m m . \quad (12.3)$$

Note that the renormalisation factors we introduced are related but not identical to the Z -factors of LSZ. The lagrangian now takes the form

$$\mathcal{L} = -\frac{1}{4} Z_A F_{\mu\nu} F^{\mu\nu} + Z_\psi \bar{\psi} (i(\not{\partial} + iZ_e Z_A^{1/2} e \not{A}) - Z_m m) \psi . \quad (12.4)$$

The key idea is as follows: Choose Z_i to be specific functions of Λ ,

$$Z_i = Z_i(\Lambda) , \quad (12.5)$$

such that $Q_j = Q_j(e, m, \Lambda, Z_i(\Lambda))$ have a well-defined (finite) limit as $\Lambda \rightarrow \infty$:

$$Q_j^\infty \equiv \lim_{\Lambda \rightarrow \infty} Q_j(e, m, \Lambda, Z_i(\Lambda)) . \quad (12.6)$$

If that is possible, our QFT is called **renormalisable**.

Note that it is highly non-trivial that this can be achieved because there are infinitely many observables, $j = 1, \dots, \infty$, while the number of Z_i 's is finite (here: $i = 1, \dots, 4$). Even if it is possible, this procedure is in general non-unique: One can always move finite factors between renormalised quantities and the Z_i . Thus, to turn the above in a unique procedure one needs so-called **Renormalisation conditions**.

For example, one may fix the positions of the poles of all propagators to the physical masses of the corresponding particles and all the residues to unity. Moreover, one may fix some specific cross section experimentally and, using its tree-level (a.k.a. LO) expression $\sigma = \sigma(e, m)$, one may then fix e . As result, all *further* observables (other cross sections, Green's or correlation functions etc.) are now unambiguously determined in terms of σ, m or, equivalently e, m . Only such *further* cross sections and correlation functions will then be predictions of the theory.

As a comment, we note that one may be slightly more economical: If one accepts that correlation functions diverge as $\Lambda \rightarrow \infty$, one may restrict oneself to introducing only the renormalisation factors $Z_e = Z_e(\Lambda)$ and $Z_m = Z_m(\Lambda)$. That is sufficient to make all 'field-normalisation-independent' observables, like cross sections and decay rates, finite.

12.2 Renormalization conditions

12.2.1 Mass and field (re)normalisation of a scalar

Though we will not need this in QED, let us start with the particularly simple case of the mass and field renormalization of the scalar field. Even more concretely, let us look at $\lambda\phi^4$ theory where, in agreement with our general approach, we now write the lagrangian as

$$\mathcal{L} = \frac{1}{2} (\partial\phi)^2 Z_\phi - \frac{1}{2} \phi^2 m^2 Z_\phi Z_m - \frac{\lambda}{4!} \phi^4 Z_\lambda Z_\phi^2 . \quad (12.7)$$

In the case of free fields, i.e. for $\lambda = 0$, we expect the Z_i 's to become unity. Therefore it is natural to write them as¹³

$$Z_\varphi = 1 + \delta Z_\varphi, \quad Z_m = 1 + \delta Z_m, \quad Z_\lambda = 1 + \delta Z_\lambda. \quad (12.8)$$

Here we think of a power series expansion of all quantities in λ , such that δZ_i collects all higher-order terms. With this, \mathcal{L} can be written as

$$\mathcal{L} = \underbrace{\frac{1}{2}(\partial\varphi)^2 - \frac{1}{2}\varphi^2 m^2}_{\text{free}} + \underbrace{\frac{1}{2}(\partial\varphi)^2 \delta Z_\varphi - \frac{1}{2}\varphi^2 m^2 (\delta Z_\varphi + \delta Z_m) - \frac{\lambda}{4!}\varphi^4 Z_\lambda Z_\varphi^2}_{\text{interactions}}. \quad (12.9)$$

Crucially, here we have used the fact that all δZ_i are of linear and higher order in λ and we have hence grouped them together with the interaction term. We will treat them as higher-order corrections in a systematic expansion in λ .

The new interaction terms introduced in this way contain two fields and hence correspond to a 2-vertex. We will denote it by

$$\text{---}\times\text{---} \quad . \quad (12.10)$$

The corresponding Feynman rule can be derived as usual. In this simple case it can also be simply read off from the above lagrangian. For example, for the correction to the mass term, the so-called mass **counterterm**, we obviously have

$$\text{---}\overset{m^2}{\times}\text{---} = im^2(\delta Z_\varphi + \delta Z_m). \quad (12.11)$$

A similar term proportional to p^2 is added to include the effect of the lagrangian term $\sim (\partial\varphi)^2 \delta Z_\varphi$. If we now calculate, for example, the leading-order self-energy correction, we find

$$-i\Pi(p^2) = \underbrace{\text{---}\times\text{---}}_{\text{counter term}} + \frac{\text{---}\bigcirc\text{---}}{\sim \Lambda^2} \quad \text{at } \mathcal{O}(\lambda). \quad (12.12)$$

At this point, we recall from Sect. 7.6 the definition

$$\text{---}\bigcirc\text{---} = -i\Pi(p^2) \quad (12.13)$$

¹³ This argument may appear to be somewhat circular concerning Z_λ which, in the absence of interactions, does not even exist. Nevertheless, it is common and useful to expand also Z_λ in this manner.

The analogue of the self-energy,

$$\mu \sim \text{[diagram: a circle with a grid pattern]}\sim \nu = i\Pi_{\mu\nu}(q), \quad (12.27)$$

is also often called **vacuum polarization** in the photon case. The term on the right-hand side can be thought of as a 4×4 -matrix which we will denote by Π^M . We have

$$\text{[diagram: a circle with a grid pattern]} = i\Pi^M; \quad (12.28)$$

$$\text{[diagram: a circle with diagonal lines]} = \frac{i}{-q^2\eta} + \frac{i}{-q^2\eta} i\Pi^M \frac{i}{-q^2\eta} + \dots = \frac{i}{-q^2\eta + \Pi^M(q)}.$$

Here η denotes the matrix $\eta^{\mu\nu}$.

We know that $\Pi^M(0) = 0$ must be maintained to keep the photon massless and we assume that our regularisation respects this. By covariance, we have

$$\Pi_{\mu\nu}(q) = \eta_{\mu\nu}A(q^2) + q_\mu q_\nu B(q^2). \quad (12.29)$$

As a crucial fact, we claim that gauge-invariance enforces $A = -B \cdot q^2$, i.e.,

$$\Pi_{\mu\nu}(q) = (\eta_{\mu\nu}q^2 - q_\mu q_\nu)\Pi(q^2). \quad (12.30)$$

For this, a general derivation will be given later on. A quick argument is as follows: Consider $2 \rightarrow n$ photon scattering:

$$\text{[diagram: a circle with a grid pattern and wavy lines]} \left. \vphantom{\text{[diagram: a circle with a grid pattern and wavy lines]}} \right\} n \text{ outgoing photons.} \quad (12.31)$$

Let us change one polarization vector by a gauge transformation:

$$\epsilon^\mu(k) \rightarrow \epsilon^\mu(k) + \alpha k^\mu. \quad (12.32)$$

Since, by gauge invariance of our theory, the amplitude should not change, we conclude that the amplitude must vanish if any index is contracted with its corresponding k^μ .

Our interest is in the special case of only two photon lines. We then expect

$$\mu \sim \text{[diagram: a circle with a grid pattern]} \sim \nu \quad \begin{matrix} k^\nu \\ k \end{matrix} = 0 \quad (12.33)$$

both fermion momenta on-shell and multiply with $\bar{u}(p')$ and $u(p)$ from left and right respectively. Then we require that in the low-energy limit, i.e. for $q \rightarrow 0$, we should recover the tree-level result: $\Gamma^\mu(p, p) = \gamma^\mu$.

It is interesting to see how many independent constraints this actually imposes. This is a non-trivial question since Γ^μ is in general a matrix.

To analyse this, let us observe that Γ^μ can be built from γ matrices and hence its index μ can only belong to a γ matrix or to one of the available independent vectors, which are p and p' . (The Levi-Civita tensor can not appear since we are dealing with a parity invariant theory.) Hence, in full generality,

$$\Gamma^\mu = \gamma^\mu \cdot A(\not{p}, \not{p}') + (p'^\mu + p^\mu) B(\not{p}, \not{p}') + (p'^\mu - p^\mu) C(\not{p}, \not{p}'). \quad (12.40)$$

Now we take advantage of the fact that we analyse Γ only as being sandwiched between $\bar{u}(p')$ and $u(p)$. Hence, we may use

$$\not{p} u(p) = m u(p'), \quad \bar{u}(p) \not{p}' = \bar{u}(p') m \quad (12.41)$$

to replace A , B , C , without loss of generality, by numbers $\times 1$. Thus, we have

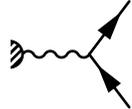
$$\Gamma^\mu = \gamma^\mu A(q^2) + (p'^\mu + p^\mu) B(q^2) + (p'^\mu - p^\mu) C(q^2), \quad (12.42)$$

where q^2 is the only kinematic invariant and hence the only possible argument of our scalar functions A, B, C .

As before, using gauge invariance we can argue that

$$\bar{u}(p') q^\mu \Gamma_\mu u(p) = 0, \quad (12.43)$$

even if $q^2 \neq 0$ since the gauge parameter within the propagator could vary and this should not affect the result. One may think here of the propagator as appearing in a more complicated diagram like



$$\quad \cdot \quad (12.44)$$

We thus learn from gauge invariance that

$$\begin{aligned} q_\mu (p'^\mu - p^\mu) = q^2 \neq 0 & \Rightarrow C = 0 \\ q_\mu (p'^\mu + p^\mu) = p'^2 - p^2 = m^2 - m^2 = 0 & \Rightarrow B \text{ unconstrained.} \end{aligned} \quad (12.45)$$

Now we have

$$\Gamma^\mu = \gamma^\mu A(q^2) + (p'^\mu + p^\mu) B(q^2). \quad (12.46)$$

Applying $\bar{u}(p')$ and $u(p)$ from left and right and using the Gordon identity

$$2m \bar{u}(p') \gamma^\mu u(p) = \bar{u}(p') (p'^\mu + p^\mu + i\sigma^{\mu\nu} q_\nu) u(p), \quad (12.47)$$

with $\sigma^{\mu\nu} \equiv \frac{i}{2}[\gamma^\mu, \gamma^\nu]$, one may trade the term proportional to $(p'^\mu + p^\mu)$ for a term proportional $\sigma^{\mu\nu}$ and a contribution to the term with γ^μ . Thus, we have

$$\Gamma^\mu = \gamma^\mu F_1(q^2) + \frac{i\sigma^{\mu\nu} q_\nu}{2m^2} F_2(q^2), \quad (12.48)$$

where $F_1(q^2)$ and $F_2(q^2)$ are known as **form factors**. We see that we are actually imposing only a single constraint: $F_1(0) = 1$.

12.3 Vacuum polarisation in dimensional regularisation

Before continuing the conceptual development, let us calculate one of the relevant, divergent loop diagrams which make the procedure of renormalisation necessary and important. While it is intuitive to think of regularising by a cutoff, to exclude high momenta, in practice a different approach (automatically respecting Poincaré and gauge invariance) is more useful. It is known as dimensional regularisation:

Let us embed our theoretical considerations in a d -dimensional space-time and try to calculate

$$i \Pi_{(1)}^{\mu\nu}(q) = \text{diagram} = (-1)(ie)^2 \int \frac{d^d k}{(2\pi)^d} \text{tr} \left[\gamma^\mu \frac{i}{\not{k} - m} \gamma^\nu \frac{i}{\not{k} + \not{q} - m} \right]. \quad (12.49)$$

Before doing so, let us explain the idea of dimensional regularization using a simpler example. Consider the log-divergent euclidean integral

$$\int^\Lambda \frac{d^4 k_E}{(k_E^2 + m^2)^2} = \Omega_3 \int_0^\Lambda \frac{d|k_E| |k_E|^3}{(|k_E|^2 + m^2)^2} \simeq \Omega_3 \ln\left(\frac{\Lambda}{m}\right). \quad (12.50)$$

In dimensional regularization with $d = 4 - \varepsilon$, the same integral gives

$$\begin{aligned} \int^\Lambda \frac{d^{4-\varepsilon} k_E}{(k_E^2 + m^2)^2} &= \Omega_{3-\varepsilon} \int_0^\infty \frac{d|k_E| |k_E|^{3-\varepsilon}}{(|k_E|^2 + m^2)^2} = \Omega_{3-\varepsilon} m^{-\varepsilon} \int_0^\infty \frac{d|k'_E| |k'_E|^{3-\varepsilon}}{(|k'_E|^2 + 1)^2} \\ &\simeq \Omega_3 m^{-\varepsilon} \int_1^\infty \frac{d|k'_E| |k'_E|^{3-\varepsilon}}{|k'_E|^4} = \Omega_3 m^{-\varepsilon} \int_1^\infty \frac{dx}{x^{1+\varepsilon}} = \Omega_3 m^{-\varepsilon} \frac{1}{\varepsilon}. \end{aligned} \quad (12.51)$$

Here, in going from the first to the second line, we used the fact that, at $\varepsilon \ll 1$, our integral is almost log-divergent. It is hence dominated by the region $|k'_E| \gg 1$. We thus make only a small error by dropping the '+1', which suppresses the regime $|k'_E| \ll 1$, and instead starting the integration at $|k_E| = 1$. In this step we also use $\Omega_{3-\varepsilon} \simeq \Omega_3$. But we keep the ε in $m^{-\varepsilon}$ since it seems awkward to change the mass dimension of the expression.

Our result demonstrates that the poles in ε track the physical log-divergence. This can be made rigorous, see e.g. Collins' book on Renormalization [8]. In the end, we always return to $d = 4$ and absorb the $1/\varepsilon$ poles in the Z_i . This is analogous to what was explained earlier using the cutoff Λ rather than ε as a regulator.

It is crucial that even for a modified number of dimensions Poincaré and gauge invariance are fully preserved. Thus we still have

$$\Pi_{\mu\nu} = (q^2 \eta_{\mu\nu} - q_\mu q_\nu) \Pi \quad (12.52)$$

and hence

$$\Pi_\mu^\mu = (q^2 d - q^2) \Pi \quad \text{and} \quad \Pi = \frac{1}{(d-1)q^2} \Pi_\mu^\mu. \quad (12.53)$$

The trace in our integral is evaluated according to

$$\begin{aligned}
\text{tr} \left[\gamma^\mu \frac{i}{\not{k} - m} \gamma^\mu \frac{i}{\not{k} + \not{q} - m} \right] &= - \frac{\text{tr} [\gamma^\mu (\not{k} + m) \gamma_\mu (\not{k} + \not{q} + m)]}{(k^2 - m^2) ((k + q)^2 - m^2)} \\
&= - \frac{\text{tr} [((2 - d)\not{k} + md) (\not{k} + \not{q} + m)]}{(k^2 - m^2) ((k + q)^2 - m^2)} \\
&= 4 \frac{(d - 2)k(k + q) - m^2 d}{(k^2 - m^2) ((k + q)^2 - m^2)}, \tag{12.54}
\end{aligned}$$

where we used the Clifford algebra in d dimensions with relations like

$$\gamma_\mu \gamma^\mu = d \cdot \mathbb{1}, \quad \gamma^\mu \not{k} \gamma_\mu = 2\not{k} - \gamma^\mu \gamma_\mu \not{k} = (2 - d)\not{k}, \quad \text{etc.} \tag{12.55}$$

Note that here the convention is to use, in spite of the changed number of dimensions, $\text{tr}(\mathbb{1}) = 4$. This is just an unimportant overall factor.

The integral in d dimensions over k of the above expression lets one expect a quadratic divergence in $d = 4$. In dimensional regularization this corresponds to a pole at $d = 2$. But at $d = 2$, the coefficient of the k^2 -term vanishes. Accordingly Π_μ^ν is not quadratically divergent in $d = 4$. That is something which is not as easily seen if one simply introduces a cutoff.

It is convenient to proceed by introducing an integration over a so-called Feynman parameter:

$$\frac{1}{AB} = \int_0^1 dx \frac{1}{(xA + (1 - x)B)^2}. \tag{12.56}$$

With this we can write

$$i\Pi_{(1)\mu}^\mu = 4e^2 \int \frac{d^d k}{(2\pi)^d} \int_0^1 dx \frac{(d - 2)k(k + q) - m^2 d}{[(1 - x)(k^2 - m^2) + x((k + q)^2 - m^2)]^2}. \tag{12.57}$$

Now we change the order of integration and substitute the integration variable according to $k' = k - xq$. Then we rename k' to k . In the denominator, we obtain

$$\left[k^2 + \underbrace{x(1 - x)q^2 - m^2}_{\equiv -\Delta} \right]^2, \tag{12.58}$$

where we introduced Δ as a convenient abbreviation. For the numerator we find

$$(d - 2) (k^2 + (1 - 2x)kq - x(1 - x)q^2) - m^2 d. \tag{12.59}$$

In this term we can drop the term $(1 - 2x)kq$ as it is odd under $k \rightarrow -k$ and the denominator is even.

Together this gives

$$i\Pi_{(1)\mu}^\mu = 4e^2 \int_0^1 dx \int \frac{d^d k}{(2\pi)^d} \frac{(d - 2) (k^2 - x(1 - x)q^2) - m^2 d}{(k^2 - \Delta)^2}. \tag{12.60}$$

In the above calculation we suppressed $i\varepsilon$ for brevity. But in fact we always had $m^2 \rightarrow m^2 - i\varepsilon$. This determines the pole structure in the k^0 -plane, cf. Fig. 18.

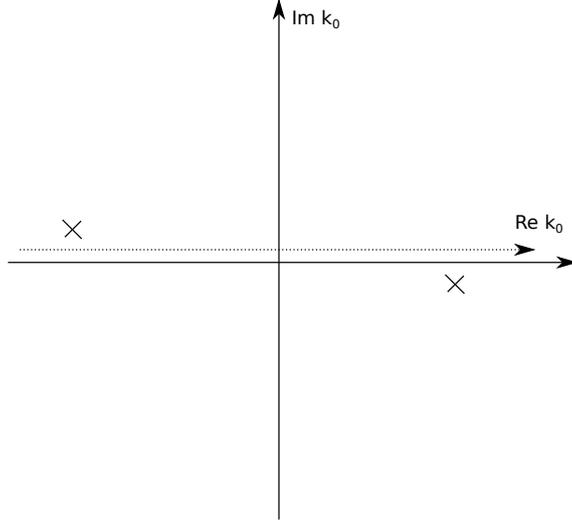


Figure 18: Pole structure in the complex k^0 plane (symbolically). In appropriate regions of q the structure is such that the integration contour can be rotated to the imaginary axis, as described in the text.

We can perform a so-called **Wick rotation**, changing the integration contour according to

$$k_0 \rightarrow ik_0 \quad , \quad dk_0 \rightarrow i dk_0 \quad , \quad k^2 = k_0^2 - \vec{k}^2 \rightarrow -k_E^2 = -(k_0^2 + \vec{k}^2). \quad (12.61)$$

The index E characterises the use of a Euclidean metric. With this we obtain

$$i\Pi_{(1)\mu}{}^\mu = 4ie^2 \int_0^1 dx \int \frac{d^d k_E}{(2\pi)^d} \frac{(d-2)(-k_E^2 + x(1-x)q^2) - m^2 d}{(k_E^2 + \Delta)^2} \quad (12.62)$$

Next we split the fraction using

$$\frac{k_E^2 + \Delta - \Delta}{(k_E^2 + \Delta)^2} = \frac{1}{k_E^2 + \Delta} - \frac{\Delta}{(k_E^2 + \Delta)^2} \quad (12.63)$$

and hence only need to calculate integrals of the type

$$\int \frac{d^d k_E}{(2\pi)^d} \frac{1}{(k_E^2 + \Delta)^n} = \int \frac{d\Omega_{d-1}}{(2\pi)^d} \int_0^\infty d|k_E| \frac{|k_E|^{d-1}}{(k_E^2 + \Delta)^n}. \quad (12.64)$$

Here the first, angular integration is well-defined for all $d > 1$. It can easily be promoted to an analytic function of d with poles using

$$\int d\Omega_{d-1} = \frac{2\pi^{d/2}}{\Gamma(d/2)}. \quad (12.65)$$

The second integral is well-defined for all $d < 2n$ and can easily be promoted to an analytic function of both d and Δ , yielding the result

$$\int_0^\infty dy \frac{y^{d-1}}{(y^2 + \Delta)^n} = \frac{\Gamma(\frac{d}{2})\Gamma(n - \frac{d}{2})}{2\Gamma(n)} \left(\frac{1}{\Delta}\right)^{n - \frac{d}{2}}. \quad (12.66)$$

Crucially for $n = 2$ and $d = 4 - \varepsilon$ we get a pole in ε :

$$\Gamma\left(n - \frac{d}{2}\right) = \Gamma\left(\frac{\varepsilon}{2}\right) = \frac{2}{\varepsilon} - \gamma + \mathcal{O}(\varepsilon), \quad (12.67)$$

where γ is the Euler-Mascheroni constant, $\gamma \approx 0.577$.

It is equally important that we do *not* get a pole in $d = 2$ from the k_E^2 term, since this would have signalled a quadratic divergence in $d = 4$. The result is proportional to

$$(d-2) \frac{\Gamma(1) \Gamma(1 - \frac{d}{2})}{2\Gamma(1)} \left(\frac{1}{\Delta}\right)^{1 - \frac{d}{2}} = -(1 - \frac{d}{2}) \Gamma(1 - \frac{d}{2}) \Delta^{\frac{\varepsilon}{2}} = -\Gamma(2 - \frac{d}{2}) \Delta^{\frac{\varepsilon}{2}}. \quad (12.68)$$

The pole also appears in $d = 4$, corresponding to a contribution to the log divergence.

Now we use equations (12.62) and (12.64), focus on the $\varepsilon \rightarrow 0$ limit and perform the x integration to get

$$\Pi_{(1)}(q^2) = \frac{1}{(d-1)q^2} \Pi_{(1)\mu}{}^\mu = -\frac{e^2}{6\pi^2\varepsilon} + \dots \quad (12.69)$$

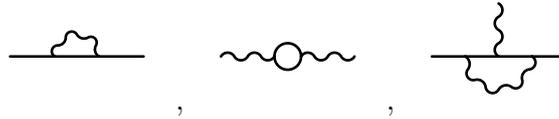
The finite terms are also easy to obtain and are important, but we will not have time to make use of them so we do not display them.

12.4 QED β -function

At this point in the course we *could* perform numerous further calculations. We could use the calculation of the last subsection to determine Z_A and the analogous self-energy calculation for the electron to fix Z_ψ and Z_m . The vertex calculation at 1-loop order would, analogously, fix Z_e . Then we could calculate two cross-sections – one to fix the coupling e in terms of experimental data, the second as a prediction of our theory.

In fact, for the first cross section, one typically picks the process $e\gamma \rightarrow e\gamma$ at $q_\gamma \rightarrow 0$. With some effort one would find that our coupling e is indeed the classical electric charge, as defined by this process.

Thus, one needs at least the diagrams



$$, \quad , \quad , \quad (12.70)$$

plus diagrams required for the actual cross-section. We do not have the time to carry this out.

However, there is a very important physical quantity (observable) which we can obtain with the help of only one diagram: **The β -function** determines the Λ -dependency of $e_0(\Lambda)$. Let us recall that our renormalized coupling e is by definition independent of Λ . As

$$e_0 = e_0(\Lambda) = Z_e(\Lambda) e, \quad (12.71)$$

we have

$$\frac{d}{d \ln(\Lambda)} e_0 = e \frac{d}{d \ln(\Lambda)} Z_e(\Lambda) = e \frac{d}{d \ln(\Lambda)} (1 + \delta Z_e(\Lambda)) \quad (12.72)$$

$$\simeq e \frac{d}{d \ln(\Lambda)} (1 + c e^2 \ln(\Lambda)) = c e^3, \quad (12.73)$$

where we used that Z has a logarithmic Λ -dependency and we are working only at leading order in e . We have defined c to be the coefficient of the 1-loop log-divergence in δZ_e .

Since $e \simeq e_0$ at leading order we can discard higher order terms on the right hand side to find

$$\frac{d}{d \ln(\Lambda)} e_0(\Lambda) = c e_0^3(\Lambda). \quad (12.74)$$

Now let us give (one possible) definition of the β -function:

$$\beta(e_0(\Lambda)) \equiv \frac{d}{d \ln(\Lambda)} e_0(\Lambda) \quad (12.75)$$

Hence equation (12.74) supplies us with the leading order β -function for the bare coupling of QED. Note that, given $\beta(e_0)$, the above differential equation (known as **renormalization group equation** or **RGE**) allows us to find e_0 for any Λ provided some boundary condition.

Why is this so-called **running** of the bare coupling of any physical relevance? To understand this, let us calculate a cross-section at a given energy \sqrt{s} at leading order:

$$\frac{d\sigma}{d\Omega} = \frac{c_1 e_0^4(\Lambda)}{s} \quad (s \gg m^2) \quad (12.76)$$

Here we used the bare coupling, which is sufficient since the higher order difference is small as long as $\ln(\Lambda/\sqrt{s})$ is not large (i.e. $\delta Z \simeq e^2 \ln(\Lambda/\sqrt{s}) \ll 1$).

Next, let us *define* a **scale-dependent physical coupling** by

$$e^4(\mu) \equiv \frac{s}{c_1} \left. \frac{d\sigma}{d\Omega} \right|_{s=\mu^2}. \quad (12.77)$$

The idea is that this coupling governs the proper strength of interactions at an energy corresponding to μ .

We see that

$$e(\mu) \simeq e_0(\Lambda) \quad \text{at} \quad \mu \lesssim \Lambda \quad (12.78)$$

and hence $e(\mu)$ obeys approximately the same RGE as e_0 :

$$\frac{d}{d \ln(\mu)} e(\mu) = \beta(e(\mu)) \quad ; \quad \beta(e) = c e^3. \quad (12.79)$$

We recall that c is defined by $\delta Z_e = c e^2 \ln(\Lambda)$.

As a cautionary remark, please note that the β -functions of the various different couplings (here bare vs. physical) agree *only at leading order*.

It is obvious that the structure $\partial_\mu + ieA_\mu$ will be unchanged under renormalization if $Z_e \sqrt{Z_A} = 1$. This relation can be shown to hold *exactly*, which is part of the **Ward-Takahashi identities**, to be discussed later on and in QFT 2.

Note that, for historical reasons, many books use a different notation:

$$Z_A = Z_3 \quad ; \quad Z_\psi = Z_2 \quad ; \quad Z_e Z_\psi \sqrt{Z_A} = Z_1. \quad (12.80)$$

Then the crucial identity we just stated takes a different form:

$$Z_e \sqrt{Z_A} = 1 \quad \Leftrightarrow \quad Z_1 = Z_2. \quad (12.81)$$

From $Z_A = Z_e^{-2}$ we can conclude that

$$c = -\frac{1}{2} \frac{1}{e^2} \frac{d}{d \ln(\Lambda)} Z_A. \quad (12.82)$$

So there is clearly a $\ln(\Lambda)$ -term in δZ_A , and δZ_A corresponds to a counterterm in \mathcal{L} :

$$\mathcal{L} \supset -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \delta Z_A = \frac{1}{2} (A_\mu \partial^2 A^\mu - A_\mu \partial^\mu \partial^\nu A_\nu) \delta Z_A. \quad (12.83)$$

This provides a counterterm vertex

$$\text{wavy line with } \times \text{ inside} = i (-\eta^{\mu\nu} p^2 + p^\mu p^\nu) \delta Z_A. \quad (12.84)$$

The complete leading-order expression for the photon self-energy now reads

$$i\Pi_{\mu\nu} = \text{wavy line with } \times \text{ inside} + \text{1-loop diagram}, \quad (12.85)$$

or in words

$$\text{self-energy} = \text{counter term} + \text{1-loop term}. \quad (12.86)$$

We also recall that

$$i\Pi_{\mu\nu} = i (q^2 \eta^{\mu\nu} - q^\mu q^\nu) \Pi(q^2) \quad (12.87)$$

and

$$i\Pi_{(1)\mu\nu} = i (q^2 \eta^{\mu\nu} - q^\mu q^\nu) \Pi_{(1)}(q^2). \quad (12.88)$$

Hence, using renormalization condition on $\Pi(q^2)$, we can conclude that

$$c = -\frac{1}{2e^2} \cdot \{ \text{coeff. of } \ln(\Lambda)\text{-term in } \Pi_{(1)}(q^2) \}. \quad (12.89)$$

Using our calculation of the vacuum-polarisation diagram, this gives us

$$c = \frac{1}{12\pi^2} \quad (12.90)$$

and thus the β -function at 1-loop level:

$$\beta(e) = \frac{e^3}{12\pi^2}. \quad (12.91)$$

Note the interesting implication this has for the interaction strength:

$$\frac{d}{d \ln(\mu)} \left(\frac{1}{e^2} \right) = -\frac{2}{e^3} \beta(e) = -\frac{2}{12\pi^2}. \quad (12.92)$$

The solution of this, illustrated in Fig. 19, shows that we find a so-called **Landau pole**. This occurs at very high energy scales (way above the Planck scale), but at the conceptual level it implies that pure QED is only effective theory for low energies. The cutoff can, in this case, not be taken all the way to infinity. This will be different, e.g., in QCD.

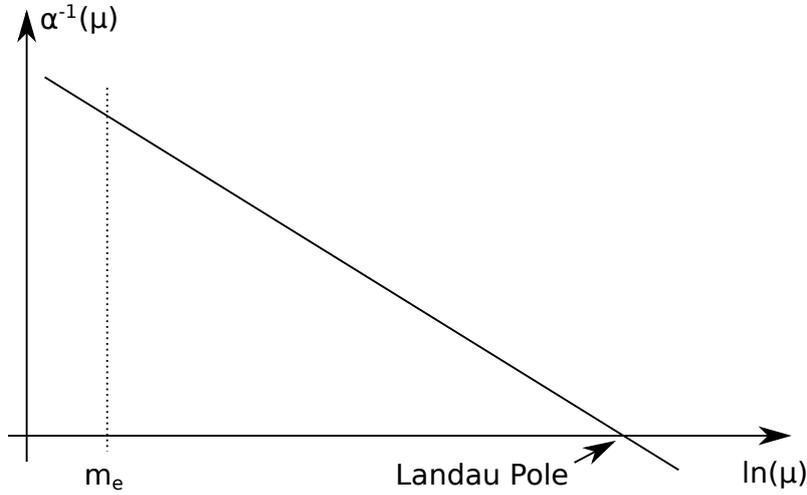


Figure 19: Inverse running coupling of QED.

12.5 The Ward-Takahashi identity

The Ward-Takahashi identity is an identity between amplitudes or Green's functions which relies on gauge invariance. We start with an illustrative example calculation: Consider a leading order three-point function with the γ -propagator amputated but the fermion propagator present, and contract this with k^μ :

$$k^\mu \cdot \left(k, \mu, \rightarrow \begin{array}{c} \text{p+k} \\ \nearrow \\ \text{p} \end{array} \right) = \frac{i}{(\not{p} + \not{k}) - m} i e \not{k} \frac{i}{\not{p} - m} = (-e) \left\{ \frac{i}{\not{p} - m} - \frac{i}{(\not{p} + \not{k}) - m} \right\} \quad (12.93)$$

The last expression follows by writing $\not{k} = (\not{p} + \not{k} - m) - (\not{p} - m)$. Pictorially this can be understood as

$$k_\mu \cdot \begin{array}{c} \xrightarrow{k} \\ \text{---} \\ \uparrow p \\ \text{---} \\ \uparrow p+k \end{array} = (-e) \left\{ \begin{array}{c} \xrightarrow{k} \\ \text{---} \\ \uparrow p \\ \text{---} \\ \uparrow p+k \end{array} - \begin{array}{c} \xrightarrow{k} \\ \text{---} \\ \uparrow p+k \\ \text{---} \\ \uparrow p \end{array} \right\} \quad (12.94)$$

In words: An external γ -line contracted with k_μ can be removed and the vertex replaced by the difference of the two propagators before and after the vertex (multiplied by $(-e)$).

We can also amputate the external fermion propagators by multiplying with

$$-((\not{p} + \not{k} - m) \{ \dots \} (\not{p} - m)) \quad (12.95)$$

which yields

$$-ek_\mu \Gamma^\mu = i \{ ((\not{p} + \not{k} - m) - (\not{p} - m)) \} \quad (12.96)$$

where in our leading order case $\Gamma^\mu = \gamma^\mu$.

Further, we can look at the actual matrix element by putting p and $p+k$ on-shell and multiplying with $\bar{u}(p+k) \dots u(p)$:

$$\bar{u}(p+k) k_\mu \Gamma^\mu u(p) = 0. \quad (12.97)$$

This is just a special case of the general gauge-invariance-based claim that $k_\mu M^\mu = 0$ for any physical amplitude with external photon.

The key interest is in the following simple generalisation of our argument:

$$\sum_{\text{all insertions of } \gamma\text{-line}} k_\mu \cdot \begin{array}{c} \xrightarrow{k} \\ \text{---} \\ \uparrow p \\ \text{---} \\ \uparrow q \end{array} = \sum_{\text{all propagators in the ferm. line}} (-e) \left\{ \begin{array}{c} \xrightarrow{k} \\ \text{---} \\ \uparrow p \\ \text{---} \\ \uparrow q \end{array} - \begin{array}{c} \xrightarrow{k} \\ \text{---} \\ \uparrow q \\ \text{---} \\ \uparrow p \end{array} \right\} \quad (12.98)$$

If we evaluate the sum, all terms except the last and the first cancel pairwise. We thus get

$$\dots = (-e) \left\{ \begin{array}{c} \xrightarrow{k} \\ \text{---} \\ \uparrow p \\ \text{---} \\ \uparrow q \end{array} - \begin{array}{c} \xrightarrow{k} \\ \text{---} \\ \uparrow q \\ \text{---} \\ \uparrow p \end{array} \right\} = (-e) \left\{ \begin{array}{c} \xrightarrow{k} \\ \text{---} \\ \uparrow p \\ \text{---} \\ \uparrow q-k \end{array} - \begin{array}{c} \xrightarrow{k} \\ \text{---} \\ \uparrow p+k \\ \text{---} \\ \uparrow q \end{array} \right\} \quad (12.99)$$

Moreover, the same argument can be made for closed loops:

$$\sum_{\text{all insertions}} k_\mu \cdot \mu \text{ --- } \vec{k} \text{ --- } \text{circle with wavy lines} = 0. \quad (12.100)$$

Here, to find the vanishing result one also needs to shift the integration variable at the end.

Combining the above statements for the open fermion line and the loop, we arrive at the **General Theorem**:

$$k_\mu \cdot \mu \text{ --- } \text{circle with wavy lines} = -e \sum_i \left\{ \text{circle with wavy lines} - \text{circle with wavy lines} \right\} \quad (12.101)$$

[Only γ -line amputated!]

For us, two corollaries are particularly important:

Corollary 1: We amputate fermion lines (multiplying with $\not{q}_i, \not{p}_i \forall i$), go on shell, and multiply with external spinors ($\bar{u}(q_i), u(p_i)$). This gives zero on the r.h. side due to factors like

$$\bar{u}(q_j) i \not{q}_j \frac{i}{\not{q}_j - \not{k}} = 0. \quad (12.102)$$

Thus, we have derived the desired result that $k_\mu \mathcal{M}^\mu = 0$ for \mathcal{M}^μ a physical amplitude.

Corollary 2: Consider the case of a 3-point-function:

$$k_\mu \cdot \mu \text{ --- } \text{circle with wavy lines} = (-e) \left(\text{circle with wavy lines} - \text{circle with wavy lines} \right) \quad (12.103)$$

Let us introduce the notation $S(p)$ for the resummed fermion propagator in momentum space:

$$S(p) = \frac{i}{\not{p} - m - \Sigma(\not{p})}. \quad (12.104)$$

The our general theorem implies

$$S(p+k)(iek_\mu\Gamma^\mu(p+k,p))S(p) = (-e)(S(p) - S(p+k)) \quad (12.105)$$

and hence

$$-iek_\mu\Gamma^\mu(p+k,p) = S^{-1}(p+k) - S^{-1}(p). \quad (12.106)$$

We can read off that the divergence in $k_\mu\Gamma^\mu$ at $k \rightarrow 0$ is the same as the divergence in Σ' . Hence it is possible to choose $Z_\psi Z_A^{1/2} Z_e = Z_\psi$ or, equivalently, $Z_1 = Z_2$. As we argued, this is the preferred choice respecting gauge invariance.

12.6 Sketch of an operator derivation of the Ward Takahashi identity

See the very end of the renormalisation chapter in the handwritten notes on the Web page of the course.

13 Non-abelian Gauge Theory and Standard Model

13.1 Non-abelian gauge theory

Remember: $\mathcal{L} = (\partial_\mu\phi)\overline{(\partial^\mu\phi)} - m^2\bar{\phi}\phi$

is invariant under $\phi(x) \rightarrow e^{i\alpha(x)}\phi(x) \Rightarrow U(1)$ -gauge-theory

Now let $\phi(x) \in V$ (vector space), such that \mathcal{L} is invariant under a group G acting on V through a representation R :

$$\phi(x) \rightarrow R(g) \cdot \phi(x), \quad \text{want } g = g(x) \quad (13.1)$$

- To be more concrete, focus on $G = SU(n)$, R is the fundamental representation

- In this case:

$$\begin{aligned} \mathcal{L} &= \overline{(\partial^\mu\phi)}_j (\partial_\mu\phi)^j - m^2\bar{\phi}_j\phi^j \\ &= (\partial_\mu\phi)^\dagger (\partial^\mu\phi) - m^2\phi^\dagger\phi \end{aligned} \quad (13.2)$$

This is obviously invariant under $\phi \rightarrow U\phi$, $U \in SU(n)$, as $U^\dagger U = \mathbb{1}$

- Let us, as in the $U(1)$ case introduce:

$$D_\mu\phi = \partial_\mu\phi + iA_\mu\phi, \quad (13.3)$$

where A_μ is a matrix and demand

$$D_\mu\phi \rightarrow UD_\mu\phi, \quad (13.4)$$

even if $U = U(x)$ (This will ensure the invariance of \mathcal{L}). This will be true if:

$$\begin{aligned} D'_\mu &\stackrel{!}{=} UD_\mu U^\dagger \\ \Leftrightarrow \mathbb{1}\partial_\mu + iA'_\mu &\stackrel{!}{=} U(\mathbb{1}\partial_\mu + iA_\mu)U^\dagger \\ \Leftrightarrow \mathbb{1}\partial_\mu + iA'_\mu &\stackrel{!}{=} \partial_\mu + U(\partial_\mu U^\dagger) + UiA_\mu U^\dagger \\ \Rightarrow A'_\mu &\stackrel{!}{=} UA_\mu U^\dagger - iU(\partial_\mu U^\dagger) \end{aligned} \quad (13.5)$$

- Let us look at the infinitesimal version $U = e^{iT}$, $T \in \mathfrak{g}$ is small.

$$L \stackrel{O}{\Rightarrow} \delta A_\mu = i[T, A_\mu] \underbrace{- \partial_\mu T}_{*} + \mathcal{O}(T^2) \quad (13.6)$$

*) We also have this in abelian gauge theory. But note: for abelian case $[.,.]$ vanishes, so the expression holds.

- We see that if $A_\mu \in \mathfrak{g}$, it will stay so! This completes our construction:

$$\begin{aligned} \mathcal{L} &= \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{matter}} \\ &= \underbrace{-\frac{1}{2g^2} \text{tr} F_{\mu\nu} F^{\mu\nu}}_{\text{gauge}} + \underbrace{(D_\mu \psi)^\dagger (D^\mu \psi) - m^2 \psi^\dagger \psi}_{\text{matter}}, \end{aligned} \quad (13.7)$$

where $F_{\mu\nu} = -i[D_\mu, D_\nu]$, $D_\mu = \partial_\mu + iA_\mu$ and $A_\mu \in \mathfrak{g}$.

- Invariance straight forwardly follows from

$$D'_\mu = U D_\mu U^\dagger \quad (13.8)$$

- In particular, $F_{\mu\nu}$ is defined as a differential operator but happens to be just a matrix. Also:

$$\begin{aligned} F'_{\mu\nu} &= U F_{\mu\nu} U^\dagger \\ \text{tr} F^2 &\text{ is invariant} \end{aligned} \quad (13.9)$$

- *Sometimes:* Convenient to choose a basis $T^a \in \mathfrak{g}$, $a = 1, \dots, \dim(G)$. Then:

$$A_\mu = A_\mu^a T^a; \quad F_{\mu\nu} = F_{\mu\nu}^a T^a \quad (13.10)$$

Consider:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + i[A_\mu, A_\nu] \quad (13.11)$$

Decompose both sides in components, use the definition of structure constants of Lie algebras $[T^a, T^b] = i f^{abc} T^c$

$$\Rightarrow F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - f^{abc} A_\mu^b A_\nu^c \quad (13.12)$$

Also if we choose a basis where $\text{tr}(T^a T^b) = \frac{1}{2} \delta^{ab}$, then

$$\text{tr}(F_{\mu\nu} F^{\mu\nu}) = \frac{1}{2} F_{\mu\nu}^a F^{\mu\nu a} \quad (13.13)$$

- Generalization to fermions:

$$\begin{aligned} \mathcal{L}_{\text{matter}} &= \bar{\psi}(i\not{D} - m)\psi \\ &= (\bar{\psi}_k)^a \left[i(\gamma^\mu)^b_a \left(\partial_\mu \delta_j^k + i A_\mu^\alpha \underbrace{(T^\alpha)_j^k}_{*} \right) - m \delta_a^b \delta_j^k \right] (\psi^j)_b \end{aligned} \quad (13.14)$$

*) possibly $R(T^\alpha)$ for other representation.

13.2 Standard Model

- We have:

$$\mathcal{L}_{\text{gauge}} = - \sum_{i=1}^3 \frac{1}{2g_i^2} \text{tr}(F_{\mu\nu}^{(i)} F^{(i)\mu\nu}), \quad (13.15)$$

where the group is $G = SU(3) \times SU(2) \times U(1)$ and $i = 1, 2, 3$ corresponds to $U(1), SU(2), SU(3)$, respectively

- Fermions come in 3 generations:

$$\mathcal{L}_{\text{matter}} = \sum_{a=1}^3 \bar{\psi}_L^a i \not{D} \psi_L^a \quad \text{all left handed} \quad (13.16)$$

where for each a , we have the fields:

$$\begin{aligned} \psi_L^a &= \{Q^a, (u^c)^a, (d^c)^a, L^a, (e^c)^a\} \\ &= \underbrace{(3, 2)_{1/3}}_* + (\bar{3}, 1)_{-4/3} + (\bar{3}, 1)_{2/3} + (1, 2)_{-1} + (1, 1)_2 \end{aligned} \quad (13.17)$$

*) 3 stands for the fundamental representation of $SU(3)$, 2 for the fundamental representation of $SU(2)$ and the index gives the charge under $U(1)$ (here: $\psi \rightarrow e^{i\alpha/3}\psi$).

The second line in 13.17 fixes all the couplings in $\mathcal{L}_{\text{matter}}$

References

- [1] Peskin/Schoeder: An Introduction to Quantum Field Theory, CRC Press, 1995.
- [2] Itzykson/Zuber: Quantum Field Theory, McGraw-Hill, 1985.
- [3] Schwartz, Quantum Field Theory and the Standard Model, Cambridge University Press, 2014.
- [4] Weinberg: The Quantum Theory of Fields, Vol. 1-3, Cambridge University Press, 2005.
- [5] Wess/Bagger: Supersymmetry and Supergravity, Princeton University Press, 1992.
- [6] Hebecker: Naturalness, String Landscape and Multiverse: A Modern Introduction with Exercises, Springer, 2021 (see also <https://arxiv.org/abs/2008.10625>).
- [7] Streater/Wightman: PCT, spin and statistics, and all that, Princeton University Press, 2000.
- [8] Collins: Renormalisation, Cambridge University Press.