

Notes on QFT

David Morrissey

January 2, 2014

Abstract

These are introductory notes on Quantum Field Theory that I developed while teaching UBC PHYS-526 in the fall 2012 and 2013. They cover much of the same material as Chapters 1–7 of Peskin and Schroeder [1], although some topics are treated in less detail here, and others in more. In writing these notes, I relied heavily on P&S, Srednicki [2], and Ryder [3]. The section on quantizing the photon is based on the treatment in Greiner and Reinhardt [19], while the section on path integrals (and Faddeev-Popov especially) follows the treatment in Pokorski [21]. Many thanks to the students in the classes for pointing out many of the typos in earlier versions.

Contents

0	Notation and Background	5
0.1	Notational Conventions	5
0.2	Background Physics	10
0.3	Useful Math	14
1	Classical Fields	17
1.1	Discretized Example: A String with Free Ends	17
1.2	Equations of Motion (Lagrangian)	20
1.3	Relativistic Actions	22
1.4	Equations of Motion (Hamiltonian)	24
1.5	Symmetries and Noether's Theorem	25
2	Quantizing the Scalar	29
2.1	Mode Expansions	29
2.2	Going Quantum	30
2.3	More on Time Dependence	33
2.4	More on Operators	34
2.5	Wave Functionals	35
3	Interacting Scalars	38
3.1	Quantizing	39
3.2	Perturbing Around the Free Theory	40
3.3	Wick's Theorem	46
3.4	A First Look at Feynman Diagrams and Rules	48

4	Feynman Rules and Scattering	53
4.1	Asymptotic States	53
4.2	Feynman Rules in Momentum Space	59
4.3	Scattering and Decays	61
5	Poincaré and Particles	66
5.1	Symmetries, Groups, and Representations	66
5.2	Symmetries in Physical Systems	70
5.3	The Poincaré Group: Lorentz plus Translations	74
5.4	Executive Summary	84
6	Fun with Fermions	85
6.1	Weyl Fermions	85
6.2	Dirac Fermions	91
7	Classical and Quantum Fermions	95
7.1	Classical Fermions	95
7.2	Quantum Fermions	100
8	Interacting Fermions	105
8.1	Perturbation Theory	105
8.2	Asymptotic States and Scattering	111
9	Quantum Electrodynamics	115
9.1	QED Lite	115
9.2	QED	118
9.3	Computing Stuff	122
10	Quantizing the Photon	124
10.1	Classical Vector Fields	124
10.2	Quantizing the Vector	128
10.3	Interacting Photons	132

11 Path Integral Quantization	134
11.1 Path Integrals for Scalar Fields	134
11.2 Path Integrals for Fermions	140
11.3 Quantizing the Photon	143
11.4 Interacting Theories	148
Bibliography	150

Chapter 0

Notation and Background

0.1 Notational Conventions

A large fraction of this course will deal with highly relativistic systems. For this reason, we will use a notation and a set of units that is geared to this situation [1, 2, 3].

0.1.1 Natural Units

We will express quantities in so-called *natural units*, defined by

$$\hbar = c = 1 , \tag{0.1}$$

where \hbar is the usual quantum mechanics thing and c is the speed of light. Since \hbar has units of energy times time, $\hbar = 1$ implies that we are measuring time in units of inverse energy. Similarly, $c = 1$ means we are measuring distance in units of time and therefore units of inverse energy as well. This simplifies dimensional analysis since now all dimensionful quantities can be expressed in units of energy. For example,

$$[E] = [P] = [M] = +1, \quad [L] = [T] = -1, \quad \left[\frac{d}{dx} \right] = +1, \tag{0.2}$$

where the square brackets denotes the energy dimension of the quantity (in natural units, of course). The specific unit we will use for energy is the electron Volt (eV), corresponding to the energy acquired by an electron passing through a potential difference of one Volt. We will also use $\text{keV} = 10^3 \text{ eV}$, $\text{MeV} = 10^6 \text{ eV}$, $\text{GeV} = 10^9 \text{ eV}$, and $\text{TeV} = 10^{12} \text{ eV}$.

To put a result back into regular units, just add powers of \hbar and c ($\sim L/T$) until you get what you want. In doing so, it's handy to remember a few things:

$$\hbar c = 1 \simeq 0.197 \text{ GeV}\cdot\text{fm} , \tag{0.3}$$

$$c = 1 \simeq 3.0 \times 10^{10} \text{ cm} , \tag{0.4}$$

$$m_p \simeq 0.938 \text{ GeV} \simeq 1.67 \times 10^{-27} \text{ kg} , \tag{0.5}$$

where $1 \text{ fm} = 10^{-13} \text{ cm}$. Despite our usage of natural units, scattering cross sections (which have units of area) will often be expressed in barns (b), with $1 \text{ b} = 10^{-24} \text{ cm}^2$. Sometimes it is also convenient to express temperatures in natural units by setting $k_B = 1$. This implies $300 \text{ K} \simeq (1/40) \text{ eV}$ (or about room temperature). Other useful mnemonics are $m_p \simeq 1 \text{ g}/N_A$ and $1 \text{ yr} \simeq \pi \times 10^7 \text{ s}$.

0.1.2 Index Notation – Vectors, Matrices, and More

Index notation will be used a lot in this class, and you'll need to be comfortable with it. As a first example, let's apply it to vectors and matrices.

An n -component vector v can be written as an $n \times 1$ matrix,

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}. \quad (0.6)$$

Clearly, we can express each of the components of v as the numbers v_i , $i = 1, 2, \dots, n$.

The same thing can be done for matrices of any size. For instance, we can write the elements of the $n \times n$ matrix M as M_{ij} :

$$M = \begin{pmatrix} M_{11} & M_{12} & \dots & M_{1n} \\ M_{21} & M_{22} & \dots & M_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n1} & M_{n2} & \dots & M_{nn} \end{pmatrix}. \quad (0.7)$$

Each entry in the matrix can be written as M_{ij} and is called a *matrix element*.

We can also use index notation to write the products of vectors and matrices. The dot product of a pair of vectors can be thought of as the matrix product of the transpose of the first with the second:

$$u \cdot v = u^t v = (u_1, u_2, \dots, u_n) \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \sum_i u_i v_i. \quad (0.8)$$

From the last equality, written in index notation, it is obvious that $u \cdot v = v \cdot u = v^t u$. Note as well that the label we use for the index that is summed over (i in this case) does not matter: $\sum_i u_i v_i = \sum_j u_j v_j$. For this reason, indices that are summed over are often called *dummy indices*.

The product of an $n \times n$ matrix M with a column vector v is itself a column vector (Mv). In components, the product is

$$(Mv)_i = \sum_j M_{ij} v_j. \quad (0.9)$$

Note here that j is a dummy index (that we can rename), while i is a *fixed index* that must match up on both sides of the equation. In this case, i labels the elements of the vector (Mv) . **Be careful not to use a fixed index to label a dummy index because you will get horribly confused and mistaken!** For example

$$(Mv)_i = \sum_k M_{ik}v_k \quad (0.10)$$

$$\neq \sum_i M_{ii}v_i . \quad (0.11)$$

The product of two matrices M and N is a matrix (MN) with elements

$$(MN)_{ij} = \sum_k M_{ik}N_{kj} = \sum_\ell M_{i\ell}N_{\ell j} . \quad (0.12)$$

Here, the i and j indices are fixed and must match up on both sides of the equation, while the k index that is summed over is a dummy index.

Index notation is also useful for expressing various matrix operations:

$$(M^t)_{ij} = M_{ji} , \quad (0.13)$$

$$(M^*)_{ij} = M_{ij}^* , \quad (0.14)$$

$$(M^\dagger)_{ij} = M_{ji}^* , \quad (0.15)$$

$$tr(M) = \sum_i M_{ii} . \quad (0.16)$$

For the special case of three-vectors, we have a cross product operation that takes a pair of vectors and makes another. It can be written in terms of indices using the *antisymmetric tensor* ϵ_{ijk} with

$$\epsilon_{123} = +1 = \epsilon_{231} = \epsilon_{312} , \quad (0.17)$$

$$\epsilon_{132} = -1 = \epsilon_{213} = \epsilon_{321} , \quad (0.18)$$

and all other entries equal to zero. This names comes from the fact that you get a factor of -1 whenever you exchange a pair of indices in ϵ . The cross product is then given by

$$(a \times b)_i = \sum_{j,k} \epsilon_{ijk} a_j b_k . \quad (0.19)$$

If you are not convinced by this, work out the components explicitly.

0.1.3 Index Notation - Relativistic

We will mostly study systems that are invariant under special relativity, meaning that the underlying equations of the system take the same form after applying Lorentz transformations (boosts and rotations). Discussing such systems is much easier if we use an index

notation appropriate to the underlying mathematical structure. Instead of writing (t, x, y, z) for a specific point in space and time, we will use

$$x^\mu = (t, x, y, z), \quad \mu = 0, 1, 2, 3. \quad (0.20)$$

We call this a (position) 4-vector, and it is useful because the components of x^μ transform linearly into each other under Lorentz transformations:

$$x^\mu \rightarrow x'^\mu = \Lambda^\mu_\nu x^\nu, \quad (0.21)$$

where Λ^μ_ν is a transformation matrix corresponding to some combination of boosts and rotations. Below, we will discuss the conditions that Λ^μ_ν must satisfy to count as a Lorentz transformation.

Any four-component object that transforms according to Eq. (0.21) is called a 4-vector. A second important example is the momentum 4-vector

$$p^\mu = (E, p^x, p^y, p^z), \quad (0.22)$$

where \vec{p} is the spatial (3-) momentum of the system and E is the energy. Recall that for a relativistic particle of mass m , we have $E = \sqrt{m^2 + \vec{p}^2}$.

For any 4-vector $a^\mu = (a^0, a^1, a^2, a^3)$ with an upper index, we define a corresponding 4-vector with a lower index by

$$a_\mu = \eta_{\mu\nu} a^\nu, \quad (0.23)$$

where the two-index object $n_{\mu\nu}$ has components

$$\eta_{\mu\nu} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (0.24)$$

In writing Eq. (0.23), we have also used Einstein's summation convention, where we implicitly sum over repeated indices.¹ In full gory detail, Eq.(0.23) is equal to

$$a_\mu = \eta_{\mu\nu} a^\nu \quad (0.25)$$

$$= \sum_{\nu=0}^4 \eta_{\mu\nu} a^\nu \quad (0.26)$$

$$= \begin{cases} a^0 & ; \quad \mu = 0 \\ -a^i & ; \quad \mu = i = 1, 2, 3 \end{cases} \quad (0.27)$$

It is really important to keep upper and lower indices distinct because it will turn out that they refer to different transformation properties under Lorentz. It is also conventional to use Greek letters for the indices of four-vectors and Roman letters for the components of three-vectors.

¹ Apparently Einstein considered this his greatest contribution to physics.

In the same way that we lower indices, we can also use the inverse of $\eta_{\mu\nu}$, called $\eta^{\mu\nu}$, to raise indices. It is not hard to see that viewed as a matrix, the components of both are the same: $\eta_{\mu\nu} = \eta^{\mu\nu}$. By definition of the inverse, we have

$$\eta^{\mu\kappa}\eta_{\kappa\nu} = \delta_{\nu}^{\mu} \quad , \quad \eta_{\mu\kappa}\eta^{\kappa\nu} = \delta_{\mu}^{\nu} \quad (0.28)$$

where $\delta_{\nu}^{\mu} = \delta_{\nu}^{\mu}$ is the 4×4 unit matrix. Raising an index with $\eta^{\mu\nu}$ is consistent with our previous definition of lowering in that if we lower an index and then raise it again, we get back what we started with:

$$a^{\mu} = \eta^{\mu\nu}a_{\nu} = \eta^{\mu\nu}(\eta_{\nu\kappa}a^{\kappa}) = \delta_{\kappa}^{\mu}a^{\kappa} = a^{\mu} . \quad (0.29)$$

These manipulations might seem trivial, but they actually have a lot of content.

A 4×4 matrix Λ_{ν}^{μ} must satisfy certain conditions for it to be a Lorentz transformation. To describe these conditions, it helps to define a 4-vector dot product by

$$a \cdot b = \eta_{\mu\nu}a^{\mu}b^{\nu} \quad (0.30)$$

$$= a_{\nu}b^{\nu} = a^{\mu}b_{\mu} \quad (0.31)$$

$$= a^0b^0 - \vec{a} \cdot \vec{b} . \quad (0.32)$$

The defining property of a Lorentz transformation is that it leave the dot product of any two vectors invariant. To be precise, if

$$a^{\mu} \rightarrow a'^{\mu} = \Lambda_{\nu}^{\mu}a^{\nu} \quad \text{and} \quad b^{\mu} \rightarrow b'^{\mu} = \Lambda_{\nu}^{\mu}b^{\nu} , \quad (0.33)$$

the matrix Λ_{ν}^{μ} must be such that

$$a' \cdot b' = a \cdot b . \quad (0.34)$$

This implies the conditions

$$\eta_{\mu\nu} \Lambda_{\lambda}^{\mu} \Lambda_{\kappa}^{\nu} = \eta_{\lambda\kappa} . \quad (0.35)$$

Equivalently, $\eta_{\nu\lambda} \eta^{\mu\kappa} \Lambda_{\kappa}^{\lambda} := \Lambda_{\nu}^{\mu} = (\Lambda^{-1})^{\mu}_{\nu}$.

Going back to our definition of the momentum 4-vector, taking its dot product with itself gives

$$p^2 := p \cdot p = E^2 - \vec{p}^2 . \quad (0.36)$$

For a relativistic particle, this is just equal to m^2 , the square of the particle (“rest”) mass which does not change under Lorentz transformations.

Using the chain rule, it is simple to show that a derivative with respect to x^{μ} transforms like a 4-vector with a lower index. As a result, we will use the simplified notation

$$\partial_{\mu} := \frac{\partial}{\partial x^{\mu}} \quad (0.37)$$

$$\partial^2 := \eta^{\mu\nu} \partial_{\mu} \partial_{\nu} = \partial_t^2 - \vec{\nabla}^2 \quad (0.38)$$

This implies

$$\partial_\mu x^\nu = \delta_\mu^\nu . \quad (0.39)$$

Note that we treat t and \vec{x} as independent variables. Thinking ahead to quantum mechanics, we will later make the identification $P^\mu = (P^0, \vec{P}) \sim i(\partial_t, -\vec{\nabla}) = i\partial^\mu$ as operators in a position-space basis.

0.2 Background Physics

Having established notation, we turn next to a quick review of the essential physics background for this course. If you are not familiar with this material, you should come speak to me as soon as possible.

0.2.1 Electromagnetism

Maxwell's equations [4]:

$$\vec{\nabla} \cdot \vec{E} = \rho \quad (0.40)$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \quad (0.41)$$

$$\vec{\nabla} \cdot \vec{B} = 0 \quad (0.42)$$

$$\vec{\nabla} \times \vec{B} = \frac{\partial \vec{E}}{\partial t} + \vec{j} \quad (0.43)$$

where ρ is the charge density and \vec{j} is the charge current. In writing the equations in this way, we have implicitly also set $\epsilon_0 = \mu_0 = 1$ (consistent with $c = 1/\sqrt{\epsilon_0\mu_0}$).

The electric and magnetic fields can be written in terms of electric and magnetic potentials:

$$\vec{E} = -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t}, \quad \vec{B} = \vec{\nabla} \times \vec{A} . \quad (0.44)$$

Down the road, we will combine the electric and magnetic potentials into a single four-vector

$$A^\mu = (\phi, \vec{A}) . \quad (0.45)$$

With these conventions, the electric potential due to a stationary point charge Q is

$$\phi = \frac{Q}{4\pi r} . \quad (0.46)$$

0.2.2 Classical Mechanics

In the Lagrangian formulation of classical mechanics, a physical system is described by a set of generalized coordinates $q_i(t)$. The time evolution of the system is determined by the action [5]:

$$S = \int_a^b dt L(q_i, \dot{q}_i) . \quad (0.47)$$

In many cases, the Lagrangian L is equal to $L = T - V$. The equations of motion that describe the time evolution of the coordinates $q_i(t)$ are derived from the principle of stationary action, namely that the physical configuration is the one that satisfies $\delta S = 0$ for *any* infinitesimal variation $q_i(t) \rightarrow q_i(t) + \delta q_i(t)$ (with $\delta q_i = 0$ for $t = a, b$). Applying this condition to the action gives the equations of motion:

$$\partial_t \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = 0 , \quad (0.48)$$

for each value of i . This gives one second-order equation for each coordinate $q_i(t)$. The solution is unique once initial values for $q_i(t)$ and $\dot{q}_i(t)$ are specified.

A second way to formulate classical mechanics is in terms of a Hamiltonian H that depends on the coordinates $q_i(t)$ and $p_i(t)$ [5]. The Hamiltonian formulation can be derived from the Lagrangian, and vice versa. Given a Lagrangian $L(q_i, \dot{q}_i)$, the generalized momenta are defined to be

$$p_i = \frac{\partial L}{\partial \dot{q}_i} . \quad (0.49)$$

Using these equations, it is possible to write each \dot{q}_i as a function of the p_i and q_i . The Hamiltonian $H(q_i, p_i)$ is then defined to be

$$H = \sum_i \dot{q}_i p_i - L , \quad (0.50)$$

where all the \dot{q}_i s on the right-hand side are to be viewed as functions of the q_i and p_i . The equations of motion in this formulation can be written in terms of Poisson brackets, defined according to

$$\{f, g\}_{PB} = \sum_j \left(\frac{\partial f}{\partial q_j} \frac{\partial g}{\partial p_j} - \frac{\partial f}{\partial p_j} \frac{\partial g}{\partial q_j} \right) . \quad (0.51)$$

With this definition, the equations of motion are

$$\dot{q}_i = \{q_i, H\}_{PB}, \quad \dot{p}_i = \{p_i, H\}_{PB} . \quad (0.52)$$

Two useful facts to keep in mind are

$$\{q_i, p_j\}_{PB} = \delta_{ij} , \quad (0.53)$$

and for any function $f(q, p, t)$

$$\frac{df}{dt} = \{f, H\}_{PB} + \frac{\partial f}{\partial t} . \quad (0.54)$$

0.2.3 Quantum Mechanics

The rules of quantum mechanics (QM) for relativistic systems are exactly the same as what you have learned in previous classes. However, the procedure we will follow to formulate a consistent quantum theory of relativistic particles will probably be new to you. In most QM courses, you are typically given a system with a Hamiltonian and asked to solve for the corresponding energy eigenvalues and eigenstates. Instead, the major challenge to be addressed in this course will be figuring out how to build the system itself and to deduce what its Hamiltonian should be. So, even though the way we approach things might seem new and unusual, keep in mind that the basic underlying rules are the same as ever.

Recall that a quantum mechanical system consists of a set of states (comprising a Hilbert space) together with the set of operators acting on them [6]. Starting from a classical system defined in terms of q_i , p_j , and H , we can construct a corresponding quantum system by elevating q_i and p_j to operators on the Hilbert space satisfying

$$[q_i, p_j] = i \delta_{ij} . \quad (0.55)$$

Notice the similarity to the Poisson brackets. The Hamiltonian H is just a function of the q_i and p_j , and is now also a well-defined operator on the Hilbert space of states (up to possible ambiguities in the order of the q_i and p_j).

Quantum states can be viewed as vectors in a (possibly infinite-dimensional) *vector space*, and observables correspond to operators on the space. The vector spaces that arise in quantum mechanics also have an *inner product* defined on them. A key result in linear algebra that we will use a lot is that the eigenstates of a Hermitian operator form a complete basis for the vector space. For example, suppose $\{|n\rangle\}$ are the eigenstates of the Hamiltonian operator H , with $H|n\rangle = E_n|n\rangle$. Completeness means that any state can be expanded in terms of them:

$$|\psi\rangle = \sum_n a_n |n\rangle , \quad (0.56)$$

where the coefficients a_n are just numbers. We can also choose to arrange the energy eigenstates to be orthonormal, $\langle m|n\rangle = \delta_{mn}$. In this case, we have

$$a_n = \langle n|\psi\rangle . \quad (0.57)$$

A convenient way to express completeness and orthonormality (which applies to any set of orthonormal basis states) is the *resolution of the identity*:

$$\mathbb{I} = \sum_n |n\rangle\langle n| , \quad (0.58)$$

where \mathbb{I} is the identity operator. Acting the identity, written in this form, on $|\psi\rangle$ then returns Eq. (0.56) with the coefficients as in Eq. (0.57).

Time evolution in quantum mechanics is determined by the Hamiltonian. There are two popular ways to describe this. The first is the Schrödinger picture in which the operators are time-independent while the state of the system evolves according to

$$|\alpha(t)\rangle = e^{-iH(t-t_0)}|\alpha(t_0)\rangle , \quad (0.59)$$

where t_0 is some reference time. Equivalently, states satisfy the Schrödinger equation,

$$\frac{d}{dt}|\alpha(t)\rangle = -iH|\alpha(t)\rangle . \quad (0.60)$$

The second standard way to describe time evolution in QM is the Heisenberg picture. Here, the states are time independent while the operators depend on time. Specifically, we have $|\alpha(t)\rangle = |\alpha(t_0)\rangle$ together with

$$\mathcal{O}(t) = e^{iH(t-t_0)}\mathcal{O}(t_0)e^{-iH(t-t_0)} . \quad (0.61)$$

This relation is equivalent to the commutation relation

$$\frac{d}{dt}\mathcal{O} = -i[\mathcal{O}, H] , \quad (0.62)$$

which should remind you of the classical Poisson bracket result.

This discussion should be very familiar from your previous courses on QM, so let's try something a bit different. Specifically, let's build a Hilbert space for a system and figure out what kinds of operators act on it.

e.g. 1. Particles in Zero Dimensions

Consider a system consisting of a box that can contain any number of identical (bosonic), non-interacting particles at rest. Let us assume further that each particle contributes an energy M to the total. A complete set of states is therefore $\{|n\rangle : n = 0, 1, 2, \dots\}$, where n labels the number of particles in the box. The total energy of state $|n\rangle$ is nM , so we have

$$H|n\rangle = nM|n\rangle .$$

At this point, we are basically done. We have a complete Hilbert space for the system and we can figure out the time evolution of any observable by applying the Schrödinger equation.

Even so, let us fiddle a bit more and look into what sorts of operators can act on the space. Just about the only thing we can do to the system is add or remove particles, so it's worth thinking about operators that would accomplish this. Let us *define* an operator that removes particles one at a time by

$$a|n\rangle = \sqrt{n}|n-1\rangle .$$

There is nothing stopping us from making such a definition; this is a perfectly well-defined operator on the Hilbert space. Taking the Hermitian conjugate of this and contracting the result with all possible kets, we also find

$$a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle .$$

From this, we can also derive

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

as well as

$$H = M a^\dagger a .$$

In fact, it can be shown that any operator whatsoever on the Hilbert space can be built up from sums and products of a and a^\dagger

All this should be familiar from the case of the harmonic oscillator. Now, however, the system does not correspond to a single point particle in a quadratic potential. Even so, we can rewrite the Hamiltonian in terms of X and P operators. For this, let us take the two linear combinations of a and a^\dagger that are Hermitian:

$$X = (a + a^\dagger), \quad P = -i(a - a^\dagger) .$$

These relations are invertible, so any operator on the Hilbert space can be constructed from X and P as well. In particular, we have

$$H = M(X^2 + P^2)/2 - M/2 . \tag{0.63}$$

Thus, our system of identical stationary bosons can be described by continuous generalized coordinate operators X and P that don't have anything to do with space or velocity.

0.3 Useful Math

The mathematics to be used in this course will consist mostly of vector calculus, together with a bit of group theory. We will get to the group theory later on, so let us just mention a few useful results in vector calculus [7].

We will write n -dimensional integrals over spacetime (one time and $(n - 1)$ spatial dimensions) as

$$\int d^n x = \int dt \int dx^1 \int dx^2 \dots \int dx^{n-1} \tag{0.64}$$

The divergence theorem generalizes to this case: for any n -dimensional volume V with boundary ∂V , we have

$$\int_V d^n x \partial_\mu f^\mu = \int_{\partial V} dA n_\mu f^\mu , \tag{0.65}$$

where n_μ is the outwardly pointing normal unit vector to the bounding surface ∂V . In most cases of interest to us, the surface integral over ∂V will be zero for one reason or another.

Multi-dimensional delta functions will appear frequently:

$$\int d^n x \delta^{(n)}(x - x') f(x) = f(x') . \tag{0.66}$$

Equivalently $\delta^{(n)}(x - x') = \prod_{\mu=0}^{n-1} \delta(x^\mu - x'^\mu)$. We also have the Fourier transform relation

$$\int d^n x e^{ik \cdot x} = (2\pi)^n \delta^{(n)}(k), \quad \int d^n k e^{ik \cdot x} = (2\pi)^n \delta^{(n)}(x). \quad (0.67)$$

Finally, recall that

$$\int dx \delta(f(x)) g(x) = \sum_a g(x_a) / |df/dx|_{x_a}, \quad (0.68)$$

where x_a refers to any value of x such that $f(x_a) = 0$.

Sometimes it will be useful to use contour integration in the complex plane. The key result is Cauchy's theorem, which is that for any integral around a closed counterclockwise contour C in the complex plane,

$$f(z_0) = \frac{1}{2\pi i} \oint_C dz \frac{f(z)}{(z - z_0)}, \quad (0.69)$$

provided $f(z)$ has no poles in the region bounded by the contour. If the contour runs clockwise, the formula picks up an overall minus sign. Differentiating both sides with respect to z_0 gives

$$\frac{d^n}{dz_0^n} f(z_0) = \frac{n!}{2\pi i} \oint_C dz \frac{f(z)}{(z - z_0)^{n+1}}. \quad (0.70)$$

This is really useful because it also applies to any deformation of the contour C that does not intersect the pole. Sometimes the stuff in the integrand (with the pole factored off) is called the *residue* of the pole. Cauchy's theorem is therefore often stated as

$$\oint_C dz f(z) = 2\pi i \times (\text{sum of residues}). \quad (0.71)$$

e.g. 2. $I = \int_0^\infty dx 1/(1 + x^2)$

We can do this integral easily using Cauchy's theorem. See Fig. 1 for an illustration of the contours C_1 and C_2 .

$$\begin{aligned} I &= \frac{1}{2} \int_{-\infty}^{\infty} dx \frac{1}{1 + x^2} \\ &= \frac{1}{2} \lim_{R \rightarrow \infty} \int_{-R}^{+R} dx \frac{1}{1 + x^2} = \lim_{R \rightarrow \infty} \int_{C_1} dz \frac{1}{1 + z^2} \\ &= \frac{1}{2} \lim_{R \rightarrow \infty} \left[\oint_{C=C_1+C_2} dz \frac{1}{(z+i)(z-i)} - \int_{C_2} dz \frac{1}{1 + z^2} \right] \\ &= \frac{1}{2} \times 2\pi i \times \frac{1}{2i} - 0 \\ &= \frac{\pi}{2}. \end{aligned}$$

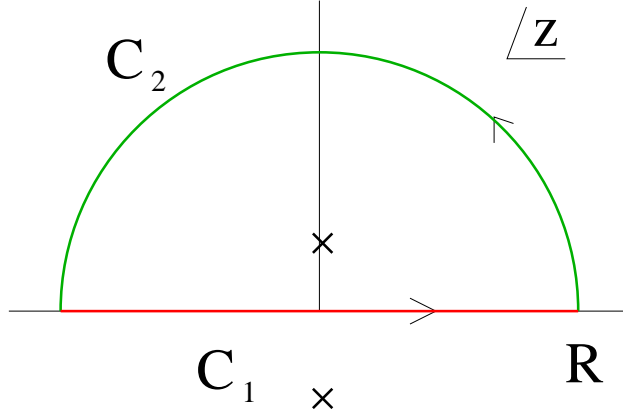


Figure 1: Contours for *e.g.* 2.

In the last line, we have used the fact that the contour integration along C_2 can be parametrized by $z = Re^{i\theta}$ and goes like

$$\lim_{R \rightarrow \infty} \int_{C_2} dz \frac{1}{1+z^2} = \lim_{R \rightarrow \infty} \int d\theta iRe^{i\theta} \frac{1}{1+R^2e^{2i\theta}} \rightarrow 0 .$$

Chapter 1

Classical Fields

A field is just a continuous system defined in time and space. You have already encountered a number of fields in your physics career.

e.g. 1. Some Fields

$T(t, \vec{x}) := T(x)$ = temperature at every point in the room.

$\phi(x)$ = electric potential at $x = (t, \vec{x})$

$\vec{E}(x)$ = electric field at $x = (t, \vec{x})$

In this example, the first and second cases are scalar fields, where the field specifies a single (real) number at every point in spacetime. The third case, the electric field, is a 3-vector field specifying a vector at every point.

To describe a system characterized by one or more fields, we must specify the equations of motion and the boundary conditions. A convenient way to encode the equations of motion is in terms of a Lagrangian. In the next example, we will illustrate how to model a continuous system as the limit of a discrete one. This example should also give you some intuition for where the various terms that usually appear in field theory Lagrangians come from. After this, we will look at how to derive the equations of motion directly, then specialize to relativistic systems, and then move on to Hamiltonians and symmetries.

1.1 Discretized Example: A String with Free Ends

Consider a string of length ℓ whose ends can move freely. The configuration of the string is described by $h(t, x)$, the height of the string element at point x and time t . Since the string has free ends, we also have the boundary conditions $\partial_x h(t, 0) = 0 = \partial_x h(t, \ell)$. We will model the system by N equally-spaced masses m connected by springs k , with each mass only allowed to move vertically. The setup is illustrated in Fig. 1.1. We will write

$$\Delta = \ell/N, \quad x_n = n\Delta, \quad (1.1)$$

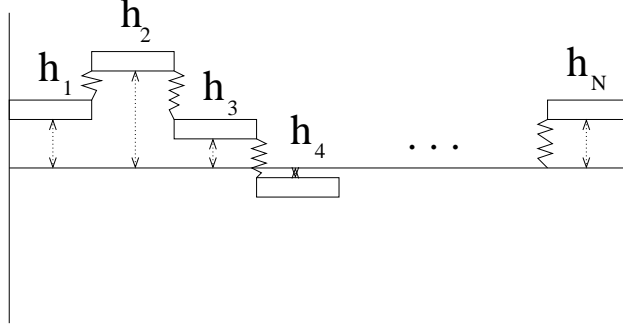


Figure 1.1: Discretized string.

and describe the height of the n -th mass by

$$h_n(t) = h(t, x_n) . \quad (1.2)$$

These are the N generalized coordinates of the discrete system.

The total kinetic energy is

$$T = \sum_{n=1}^N \frac{1}{2} m \dot{h}_n^2 . \quad (1.3)$$

The potential energy is just the sum of the relative spring displacements between each of the mass elements,

$$V = \sum_{n=1}^{N-1} \frac{1}{2} k (h_{n+1} - h_n)^2 . \quad (1.4)$$

Let us now try to connect with the continuous case by taking $N \rightarrow \infty$ (and $\Delta \rightarrow 0$) while keeping $\rho := m/\Delta$ and $\kappa := k\Delta$ fixed. The kinetic energy becomes

$$T = \sum_n \Delta \frac{1}{2} \frac{m}{\Delta} \dot{h}_n^2 \quad (1.5)$$

$$\rightarrow \int_0^\ell dx \frac{1}{2} \rho \left(\frac{\partial h}{\partial t} \right)^2 , \quad (1.6)$$

where $h = h(t, x)$, and we have used the formal definition of an integral. For the potential energy, we get

$$V = \sum_n \Delta \frac{1}{2} (k\Delta) \frac{(h_{n+1} - h_n)^2}{\Delta^2} \quad (1.7)$$

$$\rightarrow \int_0^\ell dx \frac{1}{2} \kappa \left(\frac{\partial h}{\partial x} \right)^2 . \quad (1.8)$$

We now have an action $S = \int dt (T - V)$ for the continuous system:

$$S = \int dt \int_0^\ell dx \frac{1}{2} \left[\rho \left(\frac{\partial h}{\partial t} \right)^2 - \kappa \left(\frac{\partial h}{\partial x} \right)^2 \right] \quad (1.9)$$

The integrand is sometimes called the Lagrange density.

The discretized form of the theory can also be used to deduce the equations of motion for the continuous version. Remembering that the degrees of freedom are $h_n(t) = h(t, x_n)$, we find

$$\partial_t \left(\frac{\partial L}{\partial \dot{h}_n} \right) - \frac{\partial L}{\partial h_n} = \partial_t (\Delta \rho \dot{h}_n) + \frac{\kappa}{\Delta} [(h_n - h_{n-1}) - (h_{n+1} - h_n)] \quad (1.10)$$

$$\rightarrow \Delta \rho \left(\partial_t^2 - \frac{\kappa}{\rho} \partial_x^2 \right) h(t, x) \quad (1.11)$$

Thus, the equation of motion for $h(t, x)$ is

$$\left(\partial_t^2 - \frac{\kappa}{\rho} \partial_x^2 \right) h(t, x) = 0 . \quad (1.12)$$

This is a familiar wave equation. A simple solution is

$$h(t, x) = a(k) e^{-i\omega t} e^{ipx} + (c.c.) , \quad (1.13)$$

with ω related to p by the *dispersion relation*

$$\omega = \pm p \left(\frac{\kappa}{\rho} \right)^{1/2} . \quad (1.14)$$

These solutions correspond to left- and right-moving waves with wavenumber k and equal phase and group velocities $v_p = \sqrt{\kappa/\rho} = v_g$. Applying the boundary conditions restricts the possible values of p .

A simple generalization of this example is to add the term $M^2 h^2/2$ to V . Physically, this would represent attaching each mass element to the $h = 0$ line by a spring with constant M^2 . (Note that M^2 has nothing at all to do with the mass m of the elements.) The equation of motion with this extra term becomes

$$\left(\partial_t^2 - \frac{\kappa}{\rho} \partial_x^2 + M^2 \right) h(t, x) = 0 . \quad (1.15)$$

The solution of Eq. (1.13) still works here, but with a different dispersion relation:

$$\omega = \pm \left(\frac{\kappa}{\rho} \right)^{1/2} \sqrt{p^2 + (\rho/\kappa) M^2} . \quad (1.16)$$

This should remind you of the energy of a single relativistic particle.

1.2 Equations of Motion (Lagrangian)

In the example above, we started with a continuous system, modelled it with a discrete one, and took a limit to get back to the continuous case. This works, but it is also very tedious. Instead, it is useful to apply the calculus of variations to continuous field systems directly. Suppose our system is described by the field $\phi(t, x)$ defined in one spatial dimension. In the Lagrangian formulation, we should think of the x parameter as a continuous label of generalized coordinates:

$$\phi(t, x) = q_x(t) . \quad (1.17)$$

Consider an action for the system of the form

$$S[\phi] = \int_{t_1}^{t_2} dt \left[\int_{x_1}^{x_2} dx \mathcal{L}(\phi, \partial_t \phi, \partial_x \phi) \right] . \quad (1.18)$$

Note that $L = \int dx \mathcal{L}$ here. We call \mathcal{L} the *Lagrangian density*, and we will see that the Lagrangian must take this form if we are to maintain causality.

To derive the equations of motion, we apply the principle of stationary action, namely that $\delta S = 0$ under any variation $\phi(t, x) \rightarrow \phi(t, x) + \delta\phi(t, x)$ such that $\delta\phi(t_1, x) = \delta\phi(t_2, x) = 0$. This implies

$$\delta S = \int dt \int dx \left[\frac{\partial \mathcal{L}}{\partial \phi} \delta\phi + \frac{\partial \mathcal{L}}{\partial(\partial_x \phi)} \delta(\partial_x \phi) + \frac{\partial \mathcal{L}}{\partial(\partial_t \phi)} \delta(\partial_t \phi) \right] \quad (1.19)$$

$$\begin{aligned} &= \int dt \int dx \delta\phi \left(\frac{\partial \mathcal{L}}{\partial \phi} - \partial_t \left[\frac{\partial \mathcal{L}}{\partial(\partial_t \phi)} \right] - \partial_x \left[\frac{\partial \mathcal{L}}{\partial(\partial_x \phi)} \right] \right) \quad (1.20) \\ &+ \int dt \int dx \partial_x \left[\frac{\partial \mathcal{L}}{\partial(\partial_x \phi)} \delta\phi \right] . \end{aligned}$$

In the second line, we have integrated by parts over both t and x . The boundary term in the t part vanishes due to our assumption that $\delta\phi(t_1, x) = 0 = \delta\phi(t_2, x)$, but we are still left with the x part which produces the last term in Eq. (1.20). For all the systems to be considered in this course, we will assume spatial boundary conditions such that this term vanishes ($\phi(t, x_{1,2}) = 0$ or $\partial_x \phi(t, x_{1,2}) = 0$), but you should be aware of its existence.

With this slight complication out of the way, and remembering that $\delta\phi$ is arbitrary, we find the equations of motion to be:

$$0 = \frac{\partial \mathcal{L}}{\partial \phi} - \partial_t \left[\frac{\partial \mathcal{L}}{\partial(\partial_t \phi)} \right] - \partial_x \left[\frac{\partial \mathcal{L}}{\partial(\partial_x \phi)} \right] . \quad (1.21)$$

These apply for any value of t and x in the allowed range. Taking these equations of motion and applying them to the string system considered in the previous section, it is not hard to check that they reproduce the equations of motion we found before.

The derivation given here also generalizes to systems with multiple fields defined in any number of spacetime dimensions. The Lagrangian for n spacetime dimensions ($(n-1)$ spatial

dimensions) becomes

$$L = \int d^{(n-1)}x \mathcal{L}(\phi_i, \partial_\mu \phi_i) , \quad (1.22)$$

and the equations of motion are

$$0 = \frac{\partial \mathcal{L}}{\partial \phi_i} - \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} \right] . \quad (1.23)$$

Note that we have used four-vector indices (generalized to n spacetime dimensions) to simplify the notation.

The variational principle applied to fields gives rise to the idea of functional derivatives. In n spacetime dimensions (one time, $(n-1)$ space) we define

$$\frac{\delta \phi(x_a)}{\delta \phi(x_b)} = \delta^{(n)}(x_a - x_b) = \delta(t_a - t_b) \delta^{(n-1)}(\vec{x}_a - \vec{x}_b) . \quad (1.24)$$

Note that we write $x_a = (t_a, \vec{x}_a)$ and so on. This functional derivative is just a continuous generalization of $\partial_\mu x^\nu = \delta_\mu^\nu$. We also demand that the chain rule (and product rule) hold:

$$\frac{\delta f(\phi(x))}{\delta \phi(x')} = \frac{\partial f}{\partial \phi} \frac{\delta \phi(x)}{\delta \phi(x')} . \quad (1.25)$$

The functional derivative should also commute with spacetime derivatives:

$$\frac{\delta}{\delta \phi(x')} [\partial_\mu \phi(x)] = \partial_\mu \frac{\delta \phi(x)}{\delta \phi(x')} . \quad (1.26)$$

If you don't like this, you can usually avoid it by integrating by parts.

In terms of functional derivatives (and assuming that boundary terms vanish), the equations of motion can be written as

$$0 = \frac{\delta S[\phi]}{\delta \phi(x)} . \quad (1.27)$$

This is easier to remember than varying ϕ by hand.

***e.g.* 2. Action in n spacetime dimensions**

Consider the action

$$\begin{aligned} S[\phi] &= \int dt \int d^{(n-1)}x \left[\frac{1}{2} (\partial_t \phi)^2 - \frac{\xi}{2} (\vec{\nabla} \phi)^2 - f(\phi) \right] \\ &= \int dt \int d^{(n-1)}x \left[\frac{1}{2} \phi \left(-\partial_t^2 + \xi \vec{\nabla}^2 \right) \phi - f(\phi) \right] , \end{aligned}$$

where we have integrated by parts in the second line. Taking a functional derivative gives

$$\begin{aligned}\frac{\delta S[\phi]}{\delta \phi(x')} &= \int dt \int d^{(n-1)}x \left[\left(-\partial_t^2 + \xi \vec{\nabla}^2 \right) \phi - \frac{\partial f(\phi)}{\partial \phi} \right] \delta^{(n)}(x - x') \\ &= \left(-\partial_t^2 + \xi \vec{\nabla}^2 \right) \phi(x') - \frac{\partial f(\phi)}{\partial \phi}(x') .\end{aligned}$$

This is precisely what we would have obtained from the principle of least action. Note that x in the integration is a dummy variable (in that we can rename it anything we want), while x' is fixed. The quantity in the last line is a function of x' .

1.3 Relativistic Actions

Much of this course will focus on systems that are invariant under special relativity. The basic theory we will consider for the first part of the course is that of a single real scalar field defined on all of spacetime ($t, x^i \in (-\infty, \infty)$) with action

$$S = \int d^4x \left[\frac{1}{2}(\partial_t \phi)^2 - \frac{1}{2}(\vec{\nabla} \phi)^2 - V(\phi) \right] \quad (1.28)$$

$$= \int d^4x \left[\frac{1}{2} \eta^{\mu\nu} (\partial_\mu \phi)(\partial_\nu \phi) - V(\phi) \right] , \quad (1.29)$$

where $d^4x = dt d^3x$. To cut down on writing, we will often use the standard abbreviation $\eta^{\mu\nu}(\partial_\mu \phi)(\partial_\nu \phi) = (\partial \phi)^2$. We will also assume that ϕ and $\partial_\mu \phi$ both vanish at spacetime infinity.

Compared to the previous example (*e.g.* 2) we have $\xi = 1$. Only in this case can we write the time and space derivatives as $(\partial \phi)^2$. This is essential for the theory to be Lorentz invariant, which requires that all Lorentz indices are contracted. We could have also included other fully-contracted terms such as $\phi^{23} \partial^2 \phi^5 \partial_\mu \partial_\nu \phi \partial^\mu \phi \partial^\nu \phi$. There is nothing wrong with them, but for now, let us start with the most simple possibility. The action has been also written in terms of a Lagrangian density, with terms that only depend on the single local variable x . In principle, the action could also contain *non-local* terms like

$$S \supset \int d^4x d^4y \phi(x) \phi(x - y) . \quad (1.30)$$

However, this would imply that stuff at point x can instantaneously affect stuff at point y , possibly far away. This is not consistent with causality (or locality). Thus, we will only consider local actions that can be written in terms of a Lagrangian density.

Let us begin by finding solutions to the equations of motion for

$$V(\phi) = \frac{1}{2} m^2 \phi^2 . \quad (1.31)$$

Taking a functional derivative of the action, we find

$$(-\partial^2 - m^2)\phi = 0 . \quad (1.32)$$

This is called the Klein-Gordon equation. A solution is

$$\phi(x) = a(k)e^{-ik \cdot x} \quad (1.33)$$

with $k^0 = \pm\sqrt{\vec{k}^2 + m^2}$. This looks like the energy E_k of a relativistic point particle with mass m and 3-momentum \vec{k} . The most general solution is obtained by taking linear combinations of these specific solutions and making sure the result is real (since $\phi(x)$ itself is real):

$$\phi(x) = \int \widetilde{d^3k} \left[a(\vec{k})e^{-ikx} + a^*(\vec{k})a^{ikx} \right] , \quad (1.34)$$

with $k^0 = E_k = +\sqrt{\vec{k}^2 + m^2}$ and

$$\widetilde{d^3k} = \frac{d^3k}{2E_k(2\pi)^3} . \quad (1.35)$$

This funny combination has been chosen to obtain an integration measure that is Lorentz invariant. While the basic d^3k measure is not, the following combination is:

$$\int d^4k \Theta(k^0) \delta(k^2 - m^2) = \int d^3k \frac{1}{2E_k} , \quad E_k = \sqrt{\vec{k}^2 + m^2} . \quad (1.36)$$

Another relativistic action that we will study down the road is that for electromagnetism. Let us combine the scalar and vector potentials into a single 4-vector:

$$A^\mu = (\phi, \vec{A}) . \quad (1.37)$$

In terms of them, we define the field strength tensor to be

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu . \quad (1.38)$$

Note that it is antisymmetric: $F_{\mu\nu} = -F_{\nu\mu}$. The action for electromagnetism is

$$S = \int d^4x \left(-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \right) . \quad (1.39)$$

This is invariant under Lorentz transformations because all the indices are contracted. The equations of motion derived from this action turn out to be Maxwell's Equations! (More precisely they are Maxwell's equations with no sources, $\rho = \vec{j} = 0$.)

A complication that arises in electromagnetism is that the basic variables we have used to describe the theory, A^μ , are not unique. In particular, transforming A^μ by

$$A_\mu \rightarrow A_\mu + \partial_\mu \alpha , \quad (1.40)$$

leads to the same electric and magnetic fields for any function $\alpha(x)$. Fortunately, such transformations do not change $F_{\mu\nu}$, and therefore we obtain the same equations of motion. The physical interpretation is that the A^μ are redundant variables, with different values able to describe the same physical configuration. This implies that all physical observables should be unaffected by the *gauge transformation* of Eq. (1.40).

1.4 Equations of Motion (Hamiltonian)

To connect with quantum mechanics later on, it will be useful to know the Hamiltonian formulation of our simple field theory. A sensible generalization of conjugate momenta to continuous systems is

$$\Pi(t, \vec{x}) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}(t, \vec{x})} . \quad (1.41)$$

The Hamiltonian is then

$$H = \int d^3x \mathcal{H} = \int d^3x \left[\Pi(t, \vec{x}) \dot{\phi}(t, \vec{x}) - \mathcal{L} \right] . \quad (1.42)$$

Not surprisingly, we call \mathcal{H} the Hamiltonian density. We can also generalize the concept of Poisson brackets:

$$\{f, g\}_{PB} = \int d^3x'' \left[\frac{\delta' f}{\delta' \phi(t, \vec{x}'')} \frac{\delta' g}{\delta' \Pi(t, \vec{x}'')} - \frac{\delta' g}{\delta' \phi(t, \vec{x}'')} \frac{\delta' f}{\delta' \Pi(t, \vec{x}'')} \right] \quad (1.43)$$

where the prime on the functional derivatives means that we don't touch the time variable in the sense $\delta' \phi(t, \vec{x}) / \delta' \phi(t, \vec{y}) = \delta^{(3)}(\vec{x} - \vec{y})$, and $\delta' \Pi(t, \vec{x}) / \delta' \Pi(t, \vec{y}) = \delta^{(3)}(\vec{x} - \vec{y})$. Since ϕ and Π are independent variables, we also have $\delta' \Pi(t, \vec{x}) / \delta' \phi(t, \vec{y}) = 0 = \delta' \Pi(t, \vec{x}) / \delta' \phi(t, \vec{y})$.

The Hamiltonian equations of motion are then

$$\dot{\phi}(t, \vec{x}) = \{\phi(t, \vec{x}), H\}, \quad \dot{\Pi}(t, \vec{x}) = \{\Pi(t, \vec{x}), H\} . \quad (1.44)$$

We also find that

$$\{\phi(t, \vec{x}), \phi(t, \vec{x}')\}_{PB} = 0 = \{\Pi(t, \vec{x}), \Pi(t, \vec{x}')\}_{PB} \quad (1.45)$$

$$\{\phi(t, \vec{x}), \Pi(t, \vec{x}')\}_{PB} = \delta^{(3)}(\vec{x} - \vec{x}') . \quad (1.46)$$

Note that both arguments have the same value of the time variable t .

Applying this formalism to our basic scalar field theory, the conjugate momenta are

$$\Pi(t, \vec{x}) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}(t, \vec{x})} = \dot{\phi}(t, \vec{x}) , \quad (1.47)$$

and the Hamiltonian density is

$$\mathcal{H} = \frac{1}{2} \Pi^2 + \frac{1}{2} (\vec{\nabla} \phi)^2 + V(\phi) . \quad (1.48)$$

The equation of motion for ϕ yields

$$\dot{\phi}(t, \vec{x}) = \{\phi(t, \vec{x}), H\}_{PB} \quad (1.49)$$

$$= \int d^3x' \left[\frac{\delta' \phi(t, \vec{x})}{\delta' \phi(t, \vec{x}')} \frac{\delta' H}{\delta' \Pi(t, \vec{x}')} - \frac{\delta' \phi(t, \vec{x})}{\delta' \Pi(t, \vec{x}')} \frac{\delta' H}{\delta' \phi(t, \vec{x}')} \right] \quad (1.50)$$

$$= \int d^3x' \delta^{(3)}(\vec{x} - \vec{x}') \Pi(t, \vec{x}') \quad (1.51)$$

$$= \Pi(t, \vec{x}) . \quad (1.52)$$

From the equation of motion for Π , we get

$$\dot{\Pi}(t, \vec{x}) = \{\Pi(t, \vec{x}), H\}_{PB} \quad (1.53)$$

$$= \int d^3x' \left[\frac{\delta' \Pi(t, \vec{x})}{\delta' \phi(t, \vec{x}')} \frac{\delta' H}{\delta' \Pi(t, \vec{x}')} - \frac{\delta' \Pi(t, \vec{x})}{\delta' \Pi(t, \vec{x}')} \frac{\delta' H}{\delta' \phi(t, \vec{x}')} \right] \quad (1.54)$$

$$= \int d^3x' \left[0 - \delta^{(3)}(\vec{x} - \vec{x}') \left(-\vec{\nabla}^2 \phi + \frac{\partial V}{\partial \phi} \right) \right] \quad (1.55)$$

$$= \vec{\nabla}^2 \phi - \frac{\partial V}{\partial \phi} . \quad (1.56)$$

This is equivalent to what we found in the Lagrangian formulation.

1.5 Symmetries and Noether's Theorem

Symmetries play a key role in physics, and they have interesting implications for field theories (both classical and quantum). A symmetry is a transformation of the system that leaves the physics the same. For classical theories, “the same physics” means that the equations of motion should have the same form both before and after the transformation. We will consider both discrete and continuous symmetries, but the continuous case will turn out to be more interesting.

e.g. 3. A discrete symmetry of our simple scalar theory.

Recall that the Lagrangian (density) was $\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2$. This is clearly unaffected by $\phi \rightarrow \phi' = -\phi$, and the equations of motion for ϕ' are identical to those for ϕ . On the other hand, this would not be a symmetry if V contained a term proportional to ϕ^3 .

A necessary and sufficient condition for a transformation to be a symmetry is that the action written in terms of the transformed coordinates take the same form as the original theory. There are two ways to think about this. The first is the *active picture*, where the configuration of the system is changed. For an action that is a functional of a set of fields $\{\phi_i\}$, this means that we explicitly change the fields

$$\phi_i \rightarrow \phi'_i = f_i(\phi) . \quad (1.57)$$

In turn, this changes the action according to

$$S[\phi] \rightarrow S[\phi'] := S'[\phi] , \quad (1.58)$$

where $S'[\phi]$ is some new functional of the original (untransformed) fields ϕ_i . In this active picture, a transformation is a symmetry if and only if

$$S'[\phi] = S[\phi] . \quad (1.59)$$

A standard example of an active transformation is performing a rotation on the system.

The second way to think of symmetries is in terms of the *passive picture*. Here, the idea is that we don't actually modify the system, but rather that we just use different variables to describe it. Suppose $\phi_i = f_i(\phi')$, for some new set of fields $\{\phi'_j\}$. Rewriting the action in terms of these new field variables, we find

$$S[\phi] = S[f(\phi')] = S'[\phi'] . \quad (1.60)$$

Such a transformation is a symmetry of the system if and only if

$$S'[\phi'] = S[\phi'] . \quad (1.61)$$

The standard example of a passive transformation is choosing a rotated set of axes to describe the system. Relative to the new axes, the coordinates that describe the system are different even though the system has not changed at all. You should convince yourself that both pictures are equivalent.

***e.g.* 4. A continuous symmetry for two fields.**

Consider a theory with two real fields ϕ_1 and ϕ_2 :

$$\begin{aligned} \mathcal{L} &= \frac{1}{2} [(\partial\phi_1)^2 + (\partial\phi_2)^2] - \frac{1}{2}m^2 (\phi_1^2 + \phi_2^2) \\ &= \frac{1}{2}(\partial\phi)^t(\partial\phi) - \frac{1}{2}m^2\phi^t\phi , \end{aligned}$$

where $\phi = (\phi_1, \phi_2)^t$. This is clearly invariant under the discrete transformation $\phi \rightarrow -\phi$. It is also unchanged under transformations of the form

$$\phi \rightarrow \phi' = \mathcal{O}\phi , \quad (1.62)$$

where \mathcal{O} is any 2×2 orthogonal matrix – satisfying $\mathcal{O}^t\mathcal{O} = \mathbb{I}$. These are called continuous transformations because they can be parametrized by a continuous parameter. In particular, up to a few signs, we can write any orthogonal matrix \mathcal{O} in terms of the parameter α :

$$\mathcal{O} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} . \quad (1.63)$$

Note as well that this is the active picture, and we have just argued that $S[\phi] \rightarrow S[\mathcal{O}\phi] = S'[\phi] = S[\phi]$. Alternatively, in the passive picture we would write

$$\phi = \mathcal{O}\phi'' , \quad (1.64)$$

for some new field coordinates ϕ'' . Now, $S[\phi] = S[\mathcal{O}\phi'']$ by assumption, but the fact that this is a symmetry corresponds to $S[\mathcal{O}\phi''] := S''[\phi''] = S[\phi'']$.

Continuous symmetries are especially interesting because they imply conservation laws. This relationship is called *Noether's theorem*. Let's derive it using the active picture. Consider a continuous (active) transformation described by the parameter α :

$$\phi_i(x) \rightarrow \phi'_i(x) = \phi_i(x) + \alpha \Delta \phi_i(x) , \quad (1.65)$$

where we will treat $|\alpha| \ll 1$ and work to linear order. This will change the action by changing the Lagrangian, which we assume can be written as a Lagrangian density. For this to be a symmetry we must have $S[\phi'] := S'[\phi] = S[\phi]$. Equivalently, the Lagrangian density should obey

$$\mathcal{L}(\phi') := \mathcal{L}'(\phi) = \mathcal{L}(\phi) + \alpha \partial_\mu K^\mu . \quad (1.66)$$

Note that the Lagrangian density can change by a total divergence; this will only produce a vanishing surface term in the action, and will therefore yield the same equations of motion.

Plugging the form of Eq. (1.65) into Eq. (1.66) gives

$$0 = \mathcal{L}(\phi') - \mathcal{L}(\phi) - \alpha \partial_\mu K^\mu \quad (1.67)$$

$$= \sum_i \frac{\partial \mathcal{L}}{\partial \phi_i} \alpha \Delta \phi_i + \sum_i \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} \partial_\mu (\alpha \Delta \phi_i) - \alpha \partial_\mu K^\mu \quad (1.68)$$

$$= \sum_i \left[\frac{\partial \mathcal{L}}{\partial \phi_i} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} \right] \alpha \Delta \phi_i + \alpha \partial_\mu \left[\sum_i \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} \Delta \phi_i - K^\mu \right] , \quad (1.69)$$

where we have used the product rule to move derivatives around. The terms in the first part of the last line above vanish by the equations of motion. This implies that

$$0 = \partial_\mu j^\mu, \quad \text{where} \quad j^\mu = \sum_i \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} \Delta \phi_i - K^\mu . \quad (1.70)$$

We say that the four-vector j^μ is conserved. The reason for this is that if we define the corresponding *conserved charge* (not necessarily electric charge!) by

$$Q = \int d^3x j^0 , \quad (1.71)$$

we find that

$$\partial_t Q = \int d^3x \partial_t j^0 = \int d^3x \vec{\nabla} \cdot \vec{j} = 0 . \quad (1.72)$$

Note that we get zero because everything vanishes on the boundary, by assumption. The physical interpretation of $j^\mu = (j^0, \vec{j})$ is that j^0 is a charge density and \vec{j} is a current density.

***e.g.* 5. More of our two-field example**

Applying this general result to the theory of *e.g.* 4, we find that $\mathcal{L}(\phi') = \mathcal{L}(\phi)$, and thus $K^\mu = 0$. For small rotation angles α , we also have

$$\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \rightarrow \begin{pmatrix} \phi'_1 \\ \phi'_2 \end{pmatrix} = \begin{pmatrix} 1 & -\alpha \\ \alpha & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} \phi_1 + \alpha \Delta \phi_1 \\ \phi_2 + \alpha \Delta \phi_2 \end{pmatrix} . \quad (1.73)$$

Thus, $\Delta\phi_1 = -\phi_2$ and $\Delta\phi_2 = \phi_1$. The conserved current is therefore

$$j^\mu = -(\partial^\mu\phi_1)\phi_2 + \phi_1(\partial^\mu\phi_2) . \quad (1.74)$$

It is straightforward to check that this current is indeed conserved.

A particularly important set of continuous transformations are the spacetime translations,

$$x^\lambda \rightarrow x^\lambda - a^\lambda . \quad (1.75)$$

For just about all the theories we will study in this course, these translations will be symmetries of the system. For now, let's look specifically at our simple scalar theory. It is easiest to think of the translations as an active shift of the system: $\phi(x) \rightarrow \phi'(x) = \phi(x+a)$ (with the integrals and derivatives in the action unchanged, possible because we integrate over all spacetime). For infinitesimal a^λ , we have

$$\phi'(x) = \phi(x) + a^\lambda \partial_\lambda \phi . \quad (1.76)$$

Applying this to the Lagrangian, we find

$$\mathcal{L}(\phi') = \mathcal{L}(\phi) + a^\lambda \partial_\mu (\delta_\lambda^\mu \mathcal{L}) . \quad (1.77)$$

Thus, spacetime translations are a symmetry of our theory with $K_\lambda^\mu = \delta_\lambda^\mu \mathcal{L}$. Applying our general result, the corresponding conserved currents are

$$j_\lambda^\mu = \partial^\mu \phi \partial_\lambda \phi - \delta_\lambda^\mu \mathcal{L} . \quad (1.78)$$

At this point, let us emphasize that we have just considered four different symmetries at once; each value of $\lambda = 0, 1, 2, 3$ corresponds to a different transformation.¹ In contrast, μ labels the spacetime index that always arises on the current. However, since $\eta_{\mu\nu} j_\lambda^\nu = j_{\mu\lambda} = j_{\lambda\mu}$ in this case, we can afford to be a bit careless with the indices.

For the specific case of time translations, we should take $\lambda = 0$. The corresponding charge is

$$\int d^3x j_0^0 = \int d^3x \left[\frac{1}{2} (\partial_t \phi)^2 + \frac{1}{2} (\vec{\nabla} \phi)^2 + V(\phi) \right] . \quad (1.79)$$

This is just the Hamiltonian H we derived previously (with $\Pi = \partial_t \phi$). Thus, invariance under time translations corresponds to energy conservation, $\dot{H} = 0$. Similarly, for spacial translations the related charge is

$$\int d^3x j_i^0 = \int d^3x \Pi \partial_i \phi , \quad (1.80)$$

corresponding to a conserved spatial momentum P_i . Note that the conserved spatial momentum is not the same thing as the conjugate momentum ($\Pi = \partial_t \phi$) in the Hamiltonian formulation! Given the physical interpretation of j_λ^μ , it is given a special symbol

$$j^{\mu\nu} = T^{\mu\nu} , \quad (1.81)$$

and is called the *energy-momentum tensor*. The corresponding charges are usually combined into a single conserved 4-vector, $P^\mu = \int d^3x j^{0\mu} = (H, \vec{P})$.

¹This is why we used λ instead of ν .

Chapter 2

Quantizing the Scalar

Having studied the classical theory of a real scalar field, we now apply quantum mechanics to it. Specifically, we will study the theory of the scalar $\phi(x)$ described at the classical level by the Lagrangian density

$$\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \Lambda , \quad (2.1)$$

where Λ is a constant, independent of x and ϕ . Adding this constant does not affect the equations of motion. With this Lagrangian, we find the conjugate momentum

$$\Pi(x) = \partial_t\phi(x) , \quad (2.2)$$

and the Hamiltonian

$$H = \int d^3x \left[\frac{1}{2}\Pi^2 + \frac{1}{2}(\vec{\nabla}\phi)^2 + \frac{1}{2}m^2\phi^2 + \Lambda \right] . \quad (2.3)$$

From this, we see that Λ represents a constant background energy density.

2.1 Mode Expansions

Before getting to quantization, it will be useful to look a bit more closely at the expansion of the classical field $\phi(x)$ described by the Lagrangian of Eq. (2.1) in terms of plane waves. Recall that the general solution to the classical equation of motion was

$$\phi(x) = \int \widetilde{d\vec{k}} \left[a(\vec{k})e^{-ik\cdot x} + a^*(\vec{k})e^{ik\cdot x} \right] , \quad (2.4)$$

where $k^0 = E_k = \sqrt{\vec{k}^2 + m^2}$. For the conjugate momentum $\Pi(x)$, this implies

$$\Pi(x) = \partial_t\phi(x) = -i \int \widetilde{d\vec{k}} k^0 \left[a(\vec{k})e^{-ik\cdot x} - a^*(\vec{k})e^{ik\cdot x} \right] . \quad (2.5)$$

We can also express the Hamiltonian in this way.

In formulating the quantum theory, it will be useful to solve for a and a^* in terms of ϕ and Π . To do this, recall that

$$\int d^3x e^{i\vec{p}\cdot\vec{x}} = (2\pi)^3 \delta^{(3)}(\vec{p}) . \quad (2.6)$$

Multiplying $\phi(x)$ by $e^{ip\cdot x}$, with $p^0 = E_p = \sqrt{\vec{p}^2 + m^2}$, and integrating over d^3x , we get

$$\int d^3x e^{ip\cdot x} \phi(x) = \int d^3x \int \widetilde{dk} \left[a(\vec{k}) e^{-i(k^0 - p^0)t} e^{i(\vec{k} - \vec{p})\cdot\vec{x}} + a^*(\vec{k}) e^{i(k^0 + p^0)t} e^{-i(\vec{p} + \vec{k})\cdot\vec{x}} \right] \quad (2.7)$$

$$= (2\pi)^3 \int \widetilde{dk} \left[e^{-i(k^0 - p^0)t} a(\vec{k}) \delta^{(3)}(\vec{k} - \vec{p}) + e^{i(k^0 + p^0)t} a^*(\vec{k}) \delta^{(3)}(\vec{k} + \vec{p}) \right] \quad (2.8)$$

$$= \frac{1}{2p^0} a(\vec{p}) + \frac{1}{2p^0} a^*(-\vec{p}) e^{2ip^0 t} . \quad (2.9)$$

In going from the first to the second line, we have performed the $\int d^3x$ integration to obtain the delta functions of \vec{k} and \vec{p} . Note that $\vec{k} = \pm\vec{p}$ implies $k^0 = p^0 = \sqrt{\vec{p}^2 + m^2}$. A similar calculation on Π gives

$$\int d^3x e^{ip\cdot x} \Pi(x) = (-i) \frac{1}{2} a(\vec{p}) + (i) \frac{1}{2} a^*(-\vec{p}) e^{2ip^0 t} . \quad (2.10)$$

Taking linear combinations of these two results, we find

$$a(\vec{p}) = i \int d^3x e^{ip\cdot x} (\Pi - ip^0 \phi) , \quad (2.11)$$

$$a^*(\vec{p}) = -i \int d^3x e^{-ip\cdot x} (\Pi + ip^0 \phi) . \quad (2.12)$$

It is important to note that even though $\phi(x)$ and $\Pi(x)$ both depend on t , this dependence cancels out completely in these linear combinations.

2.2 Going Quantum

We are now set to formulate the quantum theory obtained from the classical field system described by the Lagrangian of Eq. (2.1). In the classical Hamiltonian formulation of the theory, Eq. (2.3), we have the conjugate variables $\phi(x)$ and $\Pi(x)$. To define a quantum theory, we elevate $\phi(x)$ and $\Pi(x)$ to Hermitian operators on a Hilbert space. As operators, they should satisfy commutation relations analogous to the classical Poisson brackets:

$$[\phi(t, \vec{x}), \phi(t, \vec{y})] = 0 = [\Pi(t, \vec{x}), \Pi(t, \vec{y})] \quad (2.13)$$

$$[\phi(t, \vec{x}), \Pi(t, \vec{y})] = i \delta^{(3)}(\vec{x} - \vec{y}) \quad (2.14)$$

These relations are a natural generalization of the Poisson brackets we obtained in the classical theory. Note as well that these commutation relations apply specifically when ϕ

and Π are **evaluated at the same time** t .¹ This might seem a bit funny at first, but remember that we should think of \vec{x} as a label for different generalized coordinates (*i.e.* $\phi(t, \vec{x}) = q_{\vec{x}}(t)$), and that our Poisson bracket relations are defined specifically for functions of q_i and p_j evaluated at the same time value. For this reason, the relations of Eqs. (2.13,2.14) are sometimes called *equal-time commutation relations*.

At this point, it is worth going back to the mode expansions we found in Section 2.1 for the classical fields. With our quantum field operators, we can *define* time-independent operators $a(\vec{k})$ and $a^\dagger(\vec{k})$ according to

$$a(\vec{k}) = i \int d^3x e^{ik \cdot x} (\Pi - ik^0 \phi) |_{t=t_0} , \quad (2.15)$$

$$a^\dagger(\vec{k}) = -i \int d^3x e^{-ik \cdot x} (\Pi + ik^0 \phi) |_{t=t_0} , \quad (2.16)$$

where $k^0 = E_k$ and all operators on the right-hand side are to be evaluated at the arbitrary reference time t_0 . At this point, we can also invert these relations to write $\phi(t_0, \vec{x})$ and $\Pi(t_0, \vec{x})$ in terms of $a(\vec{k})$ and $a^\dagger(\vec{k})$, but only at the fixed time t_0 . We will return to the issue of time dependence later on.

With the definitions of Eqs. (2.15,2.16), the commutation relations of ϕ and Π at t_0 imply simple commutation relations for a and a^\dagger :

$$[a(\vec{k}), a(\vec{p})] = 0 = [a^\dagger(\vec{k}), a^\dagger(\vec{p})] , \quad (2.17)$$

$$[a(\vec{k}), a^\dagger(\vec{p})] = (2\pi)^3 2k^0 \delta^{(3)}(\vec{k} - \vec{p}) . \quad (2.18)$$

These relations should look familiar. Up to an overall normalization, they are equivalent to the commutation relations of the raising and lowering operators for a set of independent harmonic oscillators, $[a_i, a_j^\dagger] = i \delta_{ij}$. Using this analogy, we can build up the Hilbert space and find the energy eigenvalues.

The Hamiltonian of the system has the same form as the classical version, Eq. (2.3), but is now an operator built from the operators ϕ and Π . Rewriting it in terms of a and a^\dagger using ϕ and Π at time t_0 , we find

$$H = \frac{1}{2} \int \widetilde{d\vec{k}} k^0 [a^\dagger(\vec{k})a(\vec{k}) + a(\vec{k})a^\dagger(\vec{k})] + \Lambda V \quad (2.19)$$

$$= \int \widetilde{d\vec{k}} k^0 a^\dagger(\vec{k})a(\vec{k}) + \frac{1}{2} \int d^3k k^0 \delta^{(3)}(0) + \Lambda V , \quad (2.20)$$

where V is the volume of spacetime. The last two terms in the second line are formally infinite constants that do not affect the dynamics. We will assume that they cancel exactly. The remaining expression is just a continuous generalization of $H = \sum_{i=1}^N \omega_i a_i^\dagger a_i$, the Hamiltonian for a set of N independent oscillators.

Let's now build up the states. To do so, we make two assumptions:

¹Seriously, take note of this.

1. There exists a unique vacuum state $|0\rangle$ such that

$$a(\vec{k})|0\rangle = 0 \quad (2.21)$$

for all values of \vec{k} .

2. Any operator on the Hilbert space can be built up from $\phi(x)$ and $\Pi(x)$. Equivalently, any local operator on the Hilbert space defined at time t_0 can be constructed out of $a(\vec{k})$ and $a^\dagger(\vec{k})$.

To build the Hilbert space, we will apply powers of $a^\dagger(\vec{k})$ to the vacuum.

Consider first the state $a^\dagger(\vec{p})|0\rangle := |\vec{p}\rangle$. Applying the Hamiltonian, we find

$$H|\vec{p}\rangle = \int \widetilde{d\vec{k}} k^0 a^\dagger(\vec{k})a(\vec{k})a^\dagger(\vec{p})|0\rangle \quad (2.22)$$

$$= \int \widetilde{d\vec{k}} k^0 a^\dagger(\vec{k}) \left[a^\dagger(\vec{p})a(\vec{k}) + (2\pi)^3 2p^0 \delta^{(3)}(\vec{k} - \vec{p}) \right] |0\rangle \quad (2.23)$$

$$= 0 + p^0 |\vec{p}\rangle. \quad (2.24)$$

Therefore $|\vec{p}\rangle$ is an energy eigenstate with eigenvalue $p^0 = \sqrt{\vec{p}^2 + m^2}$. One can also show that this state is an eigenstate with eigenvalue \vec{p} of the spatial momentum operator $P^i = \int d^3x T^{0i} = - \int d^3x \Pi \partial_i \phi$ constructed from the energy momentum tensor (and generalized to quantum operators).

The physical interpretation of the state $|\vec{p}\rangle$ is that it represents a *particle* with three-momentum \vec{p} and energy $p^0 = \sqrt{\vec{p}^2 + m^2}$. In particular, the Lagrangian parameter m corresponds to the mass of the particle. This might come as a bit of a surprise. After all, we started with a continuous system, and we have ended up with discrete particle states. On the other hand, we already know that classical electromagnetism is a field theory that gives rises to particle excitations, photons, when quantized. Also, quantum mechanics is weird.

The most general state of the system is a linear combination of states of the form

$$[a^\dagger(\vec{k}_1)]^{n_1} [a^\dagger(\vec{k}_2)]^{n_2} \dots [a^\dagger(\vec{k}_N)]^{n_N} |0\rangle := |\vec{k}_1, n_1; \vec{k}_2, n_2; \dots; \vec{k}_N, n_N\rangle \quad (2.25)$$

This state has energy $E = \sum_{i=1}^N n_i k_i^0$ and three-momentum $\vec{P} = \sum_{i=1}^N n_i \vec{k}_i$. It is interpreted as a multiparticle state consisting of n_1 particles of momentum \vec{k}_1 , n_2 particles of momentum \vec{k}_2 , \dots , and n_N particles of momentum \vec{k}_N . Thus, we see that the field theory is able to describe any number of relativistic particles at once.

We can deduce three important properties of the particles described by this theory from the structure of the states. The first is that the particles have no identifying properties other than their momenta. They must therefore have spin zero. The second observation is that any number of particles can have the same momentum and there is no label to tell them apart. It follows that these particles are bosons. And third, the total energy of any multi-particle state is just the sum of the energies of the individual constituents. This implies that the particles do not interact with each other at all. For this reason, the simple quadratic theory we are studying is said to be a theory of *free particles* (or a free field theory). We will see later on that interactions between particles will emerge when we add higher-order terms to the Lagrangian.

2.3 More on Time Dependence

The operators we started with, $\phi(x) = \phi(t, \vec{x})$ and $\Pi(x) = \Pi(t, \vec{x})$, depend on time. We have also found quantum states $|\vec{k}_1, n_1; \dots; \vec{k}_N, n_N\rangle$ that are time-independent. Thus, we have been implicitly working in the Heisenberg picture of quantum mechanics where operators depend on time and states do not. Even so, there is a gap to fill. The ladder operators $a(\vec{k})$ and $a^\dagger(\vec{k})$ were constructed from ϕ and Π evaluated at the specific time t_0 , and so too was our Hamiltonian. We need to extend these results to arbitrary time values.

By assumption, time evolution in quantum mechanics is governed by the Hamiltonian. In particular, for any Heisenberg-picture operator $\mathcal{O}(t)$ we have

$$[H, \mathcal{O}] = -i\partial_t \mathcal{O} . \quad (2.26)$$

By applying this relation to successive infinitesimal time translations, one can show that it is equivalent to

$$\mathcal{O}(t) = e^{iH(t-t_0)} \mathcal{O}(t_0) e^{-iH(t-t_0)} . \quad (2.27)$$

Applying this general property to the ϕ and Π operators and using their commutation relations, we find

$$\partial_t \phi(t, \vec{x}) = i[H, \phi(t, \vec{x})] = \Pi(t, \vec{x}) , \quad (2.28)$$

$$\partial_t \Pi(t, \vec{x}) = i[H, \Pi(t, \vec{x})] = (\vec{\nabla} \phi)^2 - m^2 \phi . \quad (2.29)$$

These reproduce the classical equations of motion. However, they are now to be interpreted as relations between time-dependent operators.

We also have

$$\phi(t, \vec{x}) = e^{iH(t-t_0)} \phi(t_0, \vec{x}) e^{-iH(t-t_0)} , \quad \Pi(t, \vec{x}) = e^{iH(t-t_0)} \Pi(t_0, \vec{x}) e^{-iH(t-t_0)} . \quad (2.30)$$

Now, recall that to construct the time-independent operators $a(\vec{k})$ and $a^\dagger(\vec{k})$, we used ϕ and Π defined at the fixed reference time t_0 . We can invert Eqs. (2.15,2.16) to solve for $\phi(t_0, \vec{x})$ and $\Pi(t_0, \vec{x})$, and then use Eq. (2.30) to find the operators defined at any time at all. Taking $t_0 = 0$, the result is²

$$\phi(t, \vec{x}) = \int \widetilde{d\vec{k}} \left[a(\vec{k}) e^{-ik^0 t + i\vec{k} \cdot \vec{x}} + a^\dagger(\vec{k}) e^{ik^0 t - i\vec{k} \cdot \vec{x}} \right] , \quad (2.31)$$

and similarly for $\Pi(t, \vec{x})$. While this relation looks just like what we had for the classical fields in Section 2.1, it is now a non-trivial relationship between quantum operators.

Finally let us mention that when we defined our Hamiltonian earlier, we did so only at the specific time $t_0 = 0$. However, since $\partial_t H = [H, H] = 0$, the Hamiltonian is time-independent and our earlier definition is valid for all t . This can be verified by plugging the general expansion of Eq. (2.31) into the expression for H in terms of $\phi(t, \vec{x})$ and $\Pi(t, \vec{x})$.

²For other values of t_0 , one finds the same result but with $t \rightarrow (t - t_0)$.

2.4 More on Operators

Before moving on to study more complicated Lagrangians, it is worth spending a bit more time with the quantum field operators in the free theory. We interpreted $|\vec{p}\rangle = a^\dagger(\vec{p})|0\rangle$ as a quantum state representing a single free particle with mass m and three-momentum \vec{p} . More generally, we found that the Hilbert space \mathcal{H} can be divided according to

$$\mathcal{H} = |0\rangle \oplus \{|1\rangle\} \oplus \{|2\rangle\} \oplus \dots , \quad (2.32)$$

where $\{|1\rangle\}$ represents the subspace of one-particle states, $\{|2\rangle\}$ the subspace of two-particle states, and so on.

Applying $\phi(x)$ to the vacuum, we find

$$\phi(x)|0\rangle = \int \widetilde{dk} e^{ik \cdot x} |\vec{k}\rangle . \quad (2.33)$$

The quantum field therefore creates a linear combination of one-particle states out of the vacuum. Let us also define the operator

$$\mathbb{I}_{1p} = \int \widetilde{dk} |\vec{k}\rangle \langle \vec{k}| . \quad (2.34)$$

It is not hard to show that this operator is the identity when acting on the subspace of one-particle states: $\mathbb{I}_{1p}|\vec{p}\rangle = |\vec{p}\rangle$.

There are two especially useful operations that can be defined for products of multiple operators. The first of these is called *normal ordering*, and applies specifically to the $a(\vec{k})$ and $a^\dagger(\vec{p})$ operators. A normal-ordered operator $N\{\mathcal{O}\}$ is one in which all the a^\dagger raising operators within it are written to the left of all a lowering operators. For example,

$$N\{a(\vec{k}_1)a^\dagger(\vec{k}_2)\} = a^\dagger(\vec{k}_2)a(\vec{k}_1) , \quad (2.35)$$

$$N\{[a^\dagger(\vec{k}_1)]^{n_1}[a(\vec{k}_2)]^{n_2}[(a^\dagger(\vec{k}_3)]^{n_3}\} = [a^\dagger(\vec{k}_1)]^{n_1}[a^\dagger(\vec{k}_3)]^{n_3}[a(\vec{k}_2)]^{n_2} . \quad (2.36)$$

The nice feature of a normal-ordered operator is that it has zero expectation value when sandwiched between $\langle 0|$ and $|0\rangle$ unless it is the identity. Furthermore, using the raising and lowering commutation relations, any operator can be written uniquely as a sum of normal-ordered operators, each containing an equal or smaller number of a and a^\dagger factors. Note that sometimes normal ordering is denoted by $:\mathcal{O}:$.

The second useful operation that applies to products of ϕ and Π operators is *time ordering*. For a pair of fields, the time-ordered product is

$$T\{\phi(t_1, \vec{x}_1)\phi(t_2, \vec{x}_2)\} = \Theta(t_1 - t_2)\phi(t_1, \vec{x}_1)\phi(t_2, \vec{x}_2) + \Theta(t_2 - t_1)\phi(t_2, \vec{x}_2)\phi(t_1, \vec{x}_1) , \quad (2.37)$$

where $\Theta(t)$ is the usual step function ($\Theta(t) = 1$ for $t > 0$, $\Theta(t) = 0$ for $t < 0$). Put another way, a time-ordered product of a pair of fields means that the field with the larger time is written to the left. More generally, the time-ordered product of multiple fields is the one where they are written in order of decreasing time values, from left to right.

The vacuum expectation value of the time-ordered product of a pair of fields will be of particular importance to us. Assuming $t_1 > t_2$, we have

$$\langle 0|T\{\phi(x_1)\phi(x_2)\}|0\rangle = \int \widetilde{dk} \int \widetilde{dp} \langle 0|a(\vec{k})a^\dagger(\vec{p})|0\rangle e^{-ik\cdot x_1} e^{ip\cdot x_2} \quad (2.38)$$

$$= \int \frac{d^3k}{(2\pi)^3} \frac{1}{2k^0} e^{-ik\cdot(x_1-x_2)} \quad (2.39)$$

$$= \int \frac{d^3k}{(2\pi)^3} \int_{-\infty}^{\infty} \frac{dx}{(2\pi)} \frac{i}{x^2 - \vec{k}^2 - m^2 + i\epsilon} e^{-ix(t_1-t_2) + i\vec{k}\cdot(\vec{x}_1-\vec{x}_2)} \quad (2.40)$$

In the last line, we have rewritten the preceding line using the integral relation you derived in hw-00. Note that the $i\epsilon$ picks out the pole at $x = k^0 = \sqrt{\vec{k}^2 + m^2}$ for $t_1 > t_2$. For $t_1 < t_2$, it turns out that we obtain the same final result, but now the $i\epsilon$ selects the pole at $x = -k^0$ in the x integration. This expression is so important we give it a special name: the *Feynman propagator* D_F . It can be rewritten in the form

$$D_F(x-y) = \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2 + i\epsilon} e^{-ip\cdot(x-y)}, \quad (2.41)$$

where now p^0 is a free variable unrelated to $\sqrt{\vec{p}^2 + m^2}$.

We will see that the Feynman propagator describes a particle (or antiparticle) propagating from one point to another, possibly before and after scattering with other stuff. It also has a couple of useful properties. The first is that D_F is really simple if we Fourier transform it:

$$\int d^4x e^{ik\cdot x} D_F(x) = \frac{i}{k^2 - m^2 + i\epsilon} := \widetilde{D}_F(k), \quad (2.42)$$

with k^0 still independent of $\sqrt{\vec{k}^2 + m^2}$. The second special property of the Feynman propagator is that it almost satisfies the Klein-Gordon equation:

$$(\partial^2 + m^2)D_F(x-y) = -i\delta^{(4)}(x-y). \quad (2.43)$$

Because of this, D_F is said to be a Green's function of the Klein-Gordon equation.

2.5 Wave Functionals

We have constructed our quantum field theory in terms of states and operators in the Heisenberg picture. It can also be formulated in a way that is analogous to the wavefunction formulation of single-particle quantum mechanics in the Schrödinger picture. This formulation isn't used very much, but it's fun to see. More detailed accounts can be found in Refs. [8, 9].

Recall that for a single particle quantum system (in one dimension), we can define a

position operator \hat{x} and eigenstates such that³

$$\hat{x}|x'\rangle = x'|x'\rangle \quad , \quad \langle x'|x''\rangle = \delta(x' - x'') \quad , \quad \mathbb{I} = \int dx' |x'\rangle\langle x'| \quad . \quad (2.44)$$

The wavefunction of a state $|\Psi\rangle$ is then given by

$$\Psi(x) = \langle x|\Psi\rangle \quad . \quad (2.45)$$

This generalizes straightforwardly to multiple spatial dimensions.

The analogous set of states in a quantum theory are

$$\hat{\phi}(0, \vec{x}) |\phi'\rangle = \phi'(\vec{x}) |\phi'\rangle \quad . \quad (2.46)$$

Here, $\hat{\phi}(0, \vec{x})$ is the field operator at $t = 0$ while $\phi'(\vec{x})$ is just a function that specifies the configuration of the field system at each point in space. Since $\hat{\phi}(0, \vec{x})$ is a Hermitian operator, these states form a complete set for the system, and the different eigenvalues are mutually orthogonal. Completeness translates into

$$\mathbb{I} = \int [\mathcal{D}\phi'] |\phi'\rangle\langle\phi'| \quad . \quad (2.47)$$

Here, $\int [\mathcal{D}\phi']$ denotes a functional integral, in which all possible field configurations $\phi'(\vec{x})$ of the system are summed over. The statement of orthogonality takes the form

$$\langle\phi'|\phi''\rangle = \delta[\phi' - \phi''] \quad , \quad (2.48)$$

where $\delta[f]$ is the delta functional, equal to zero unless $f = 0$ and infinite (in a controlled way) otherwise.⁴ The normalization of the delta functional is such that

$$\int [\mathcal{D}\phi'] \delta[\phi' - \phi''] F[\phi'] = F[\phi''] \quad (2.49)$$

for any functional F .

Given a quantum state $|\Psi\rangle$ of the theory, the corresponding *wavefunctional* is

$$\Psi[\phi'] = \langle\phi'|\Psi\rangle \quad . \quad (2.50)$$

Alternatively, $\Psi[\phi']$ is just the expansion coefficient when expanding this state in the $\{|\phi'\rangle\}$ basis. We can also expand operators in this basis, and represent them as operators on wavefunctionals. The results for $\hat{\phi}(\vec{x})$ and $\hat{\Pi}(\vec{x})$ are

$$\hat{\phi}(0, \vec{x}) \rightarrow \phi(\vec{x}), \quad \hat{\Pi}(0, \vec{x}) \rightarrow -i \frac{\delta}{\delta\phi(\vec{x})} \quad . \quad (2.51)$$

³ Note that I've put a hat on the operator to distinguish it from the eigenvalue. I'll do so in this section for clarity, but nowhere else.

⁴ Recall that a functional is an object that takes in a function and outputs a single number. A familiar example is the classical action $S[\phi]$.

The ϕ case should be obvious, while for Π , note that for any functional $\Psi[\Phi]$

$$\left[\phi(\vec{x}), -i \frac{\delta}{\delta \phi(\vec{y}')} \right] \Psi[\phi'] = i \delta^{(3)}(\vec{x} - \vec{y}') \Psi[\phi'] . \quad (2.52)$$

This is precisely the commutator we want.

With this functional representation in hand, we can now expand the general Schrödinger equation in the $\{|\phi'\rangle\}$ basis to obtain a Schrödinger equation for wavefunctionals. We have

$$i \frac{d}{dt} \Psi[\phi] = \frac{1}{2} \int d^3x \left(\left[-i \frac{\delta}{\delta \phi(\vec{x})} \right]^2 + (\vec{\nabla} \phi)^2 + m^2 \phi^2 \right) \Psi[\phi] . \quad (2.53)$$

This is nothing more than a continuous generalization of the simple harmonic oscillator.

Chapter 3

Interacting Scalars

We have found that quantizing a scalar field theory with $V(\phi) = \frac{1}{2}m^2\phi^2 + \Lambda$ produces a theory of identical free particles with mass m , in which the individual particles do not interact with each other at all. Due to the simplicity of this theory, we were able to find all the energy eigenstates and eigenvalues. On the other hand, this simplicity also means that the theory is not very interesting.

In this note, we will look at more complicated scalar theories with higher-order terms in the potential. Specifically, we will investigate the Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \Lambda - \Delta V(\phi) . \quad (3.1)$$

In many cases, we will take

$$\Delta V = \frac{\lambda}{4!}\phi^4 , \quad (3.2)$$

where λ is a constant parameter. For this choice of ΔV , the classical equation of motion is

$$(\partial^2 + m^2)\phi = -\frac{\lambda}{3!}\phi^3 . \quad (3.3)$$

This is a non-linear partial differential equation. It is much harder to solve than a linear equation because it is no longer always true that a linear combination of two individual solutions is also solution. In general, we don't know how to solve such equations analytically.

Turning to the Hamiltonian formulation of the classical theory, we find the conjugate momentum

$$\Pi(x) = \partial_t\phi(x) , \quad (3.4)$$

and the Hamiltonian

$$H = \int d^3x \left[\frac{1}{2}\Pi^2 + \frac{1}{2}(\vec{\nabla}\phi)^2 + \frac{1}{2}m^2\phi^2 + \Lambda \right] + \int d^3x \Delta V(\phi) \quad (3.5)$$

$$:= \bar{H}_0 + \Delta H , \quad (3.6)$$

where $\Delta H = \int d^3x \Delta V(\phi)$.

3.1 Quantizing

To quantize this theory, we proceed just like before by elevating $\phi(x)$ and $\Pi(x)$ to operators on a Hilbert space and imposing canonical commutation relations on them at equal times. Also as before, we can define $a(\vec{k})$ and $a^\dagger(\vec{k})$ operators at the fixed reference time $t_0 = 0$ according to

$$a(\vec{k}) = i \int d^3x e^{ik \cdot x} (\Pi - ik^0 \phi)|_{t=0}, \quad (3.7)$$

$$a^\dagger(\vec{k}) = -i \int d^3x e^{-ik \cdot x} (\Pi + ik^0 \phi)|_{t=0}, \quad (3.8)$$

where $k^0 = E_k = \sqrt{\vec{k}^2 + m^2}$. Using the commutators of ϕ and Π at $t = 0$, the commutators of the a and a^\dagger operators are

$$[a(\vec{k}), a^\dagger(\vec{p})] = (2\pi)^3 2k^0 \delta^{(3)}(\vec{k} - \vec{p}), \quad [a(\vec{k}), a(\vec{p})] = 0 = [a^\dagger(\vec{k}), a^\dagger(\vec{p})]. \quad (3.9)$$

Building up the Hamiltonian at $t = 0$ in terms of the modes, we find

$$H = \int \widetilde{d\vec{k}} k^0 a^\dagger(\vec{k}) a(\vec{k}) + \int d^3x \Delta V(\phi) \quad (3.10)$$

$$= H_0 + \Delta H(t = 0). \quad (3.11)$$

Even though we have constructed the Hamiltonian from quantities defined at $t = 0$, it is still time-independent,¹ and so this expression holds for any time: $H(t) = H(0)$. We can also expand $\Delta H(0)$ in terms of the ladder operators, but the resulting expression will typically be complicated. For $\Delta H = \lambda \phi^4/4!$, it will involve products of four a and a^\dagger factors.

Using the a and a^\dagger operators, we can also build the Hilbert space. As before we assume there exists a state $|0\rangle$ that is annihilated by all the $a(\vec{k})$. In particular, $H_0|0\rangle = 0$. A complete basis of eigenstates of the Hermitian operator H_0 can then be formed by applying powers of $a^\dagger(\vec{p}_i)$ to $|0\rangle$. This is good – we have now quantized the theory by specifying the degrees of freedom and finding a basis for the corresponding Hilbert space.

Now for the bad news. For just about any ΔV that isn't linear or quadratic in ϕ , we do not know how to find the exact time evolution of the system. This difficulty is easiest to see in the Schrödinger picture, where operators are time-independent and the states evolve in time. If we try to identify $|0\rangle$ with a zero-particle state at $t = 0$, as time goes by it will evolve to $|0(t)\rangle = e^{-iHt}|0\rangle$. Since $H = (H_0 + \Delta H)$, and the ΔH piece contains stuff like $a[a^\dagger]^3$ (for $\lambda \phi^4/4!$), the time-evolved “vacuum” state will contain multi-particle components.

¹This is just one of the basic assumptions of quantum mechanics for closed systems.

We can also see the challenge in the Heisenberg picture. Starting from $a(\vec{k})$ defined at $t = 0$, we can evolve it forward in time in the usual way:

$$a(t, \vec{k}) := e^{iHt} a(\vec{k}) e^{-iHt} , \quad (3.12)$$

and similarly for a^\dagger . In the free theory, we had the very special result that $a(t, \vec{k}) = e^{-ik^0 t} a(\vec{k})$; time evolution only changed the mode operators by a phase. Unfortunately, this simple result no longer holds in the interacting theory, and $a(t, \vec{k})$ will be a complicated linear combination of products of multiple $a(\vec{k})$ and $a^\dagger(\vec{k})$ operators. In general, we do not know how to find the energy eigenstates of the full Hamiltonian when $\Delta V \neq 0$. Whatever shall we do?

3.2 Perturbing Around the Free Theory

The way out of this pit of despair is to give up on trying to compute the energy eigenstates exactly, and content ourselves with approximations to the truth. To do this, we will start with the free theory ($\Delta H = 0$), and expand around it in powers of ΔH . Such an expansion will be useful if ΔH is in some sense small relative to H_0 . We will also focus on the relatively easy problem of scattering, where particles come in from spatial infinity at $t \rightarrow -\infty$, scatter with each other, and travel off to spatial infinity at $t \rightarrow +\infty$. This lends itself well to expanding around the free theory because particles separated by large distances, as they are before and after scattering, are expected to behave like free particles. On a more technical side, we will show in `notes-04` that the only quantities needed to compute scattering amplitudes are the *n-point* functions $\langle \Omega | T \{ \phi(x_1) \phi(x_2) \dots \phi(x_n) \} | \Omega \rangle$, where $|\Omega\rangle$ is the vacuum state of the full theory.

The first step in perturbation theory is to construct a set of eigenstates of H_0 . Note that even though $H_0 \neq H$, it is still a Hermitian operator and its eigenstates will therefore be a complete and orthonormalizable set. In fact, we have already done this: the lowest state is $|0\rangle$ all the higher states that can be built from it by applying factors of $a^\dagger(\vec{p})$.

In expanding around these H_0 eigenstates, we will make two physically reasonable assumptions. They can be justified when the interaction ΔH is sufficiently small.² They are:

1. There exists a unique, normalizable, lowest-energy ground state $|\Omega\rangle$ of the full Hamiltonian H with no 3-momentum (*i.e.* $\vec{P}|\Omega\rangle = 0$) and non-zero overlap with $|0\rangle$:

$$\langle 0 | \Omega \rangle \neq 0 . \quad (3.13)$$

Note that in general, $|0\rangle$ is different from $|\Omega\rangle$.

2. The next state in the spectrum is an isolated one-particle state with momentum \vec{p} and energy $E = \sqrt{\vec{p}^2 + M^2}$ for some mass M (possibly different from m). By isolated, we mean that there is a non-zero energy gap above $|\Omega\rangle$, and another energy gap between the one-particle state (in the Lorentz frame with $\vec{p} = \vec{0}$) and the next set of states.

²However, they can be explicitly violated at *strong coupling*, when ΔH cannot be treated as small.

3. The field has vacuum expectation value of zero, $\langle \Omega | \phi(0) | \Omega \rangle = 0$.

The first assumption means that the interaction does not change the vacuum of the theory in too radical a way. The second and third imply that we will still be able to associate the field $\phi(x)$ with a specific particle species.

3.2.1 The Interaction Picture

To expand the full theory around the free theory, let us begin by writing $\phi(0, \vec{x})$ in terms of the ladder operators by inverting Eqs. (3.7,3.8):

$$\phi(0, \vec{x}) = \int \widetilde{d\vec{k}} \left[a(\vec{k}) e^{i\vec{k}\cdot\vec{x}} + a^\dagger(\vec{k}) e^{-i\vec{k}\cdot\vec{x}} \right]. \quad (3.14)$$

To evolve the field to later times, we use the full Hamiltonian:

$$\phi(t, \vec{x}) = e^{iHt} \phi(0, \vec{x}) e^{-iHt}. \quad (3.15)$$

Unfortunately, we don't know how to simplify this because we don't know how to move H through the $a(\vec{k})$ and $a^\dagger(\vec{k})$ operators that make up $\phi(0, \vec{x})$.

In the face of this challenge, let us do something easier and define the time-dependent *interaction picture* field by

$$\phi_I(t, \vec{x}) := e^{iH_0 t} \phi(0, \vec{x}) e^{-iH_0 t} \quad (3.16)$$

$$= \int \widetilde{d\vec{k}} \left[a(\vec{k}) e^{-ik\cdot x} + a^\dagger(\vec{k}) e^{ik\cdot x} \right], \quad (3.17)$$

where H_0 is the free Hamiltonian at $t = 0$. Here, the time evolution is simple because we do know how to commute H_0 with a and a^\dagger . We can also generalize this definition to any other local operator: $\mathcal{O}_I(t) := e^{iH_0 t} \mathcal{O}(0) e^{-iH_0 t}$.

Before moving on, let us briefly take note of the time dependence of the full Hamiltonian and its pieces. We have

$$H(t) = e^{iHt} H(0) e^{-iHt} = H(0) \quad (3.18)$$

$$= e^{iHt} H_0 e^{-iHt} + e^{iHt} \Delta H(0) e^{-iHt} \quad (3.19)$$

$$:= \tilde{H}_0(t) + \Delta H(t). \quad (3.20)$$

While the full Hamiltonian is time-independent (since it commutes with itself), the terms within it need not be. We have defined here a time-dependent version $\tilde{H}_0(t)$ of the free Hamiltonian such that $H_0 = \tilde{H}_0(0)$. The tilde is to distinguish it from H_0 , which is given by the explicit expression in Eq. (3.10).

The interaction picture is useful because it factors out the time evolution due to the free Hamiltonian. We can relate any Heisenberg-picture operator to the interaction picture version by

$$\mathcal{O}(t) = U_I^\dagger(t) \mathcal{O}_I(t) U_I(t), \quad (3.21)$$

where

$$U_I(t) = e^{iH_0 t} e^{-iHt} . \quad (3.22)$$

Differentiating with respect to t , this implies

$$i\partial_t U_I(t) = \Delta H_I(t) U_I(t) , \quad (3.23)$$

where $\Delta H_I(t) = e^{iH_0 t} \Delta H(0) e^{-iH_0 t}$ is the interaction Hamiltonian in the interaction picture. Note that if $\Delta H = \int d^3x \lambda^{(m,n)} \Pi^m \phi^n$, we have $\Delta H_I = \int d^3x \lambda^{(m,n)} \Pi_I^m \phi_I^n$.

The solution to Eq. (3.23) for $t > 0$ is derived in detail in Peskin & Schroeder [1], and is given by *Dyson's equation*:

$$U_I(t) = T \left\{ \exp \left[-i \int_0^t dt' \Delta H_I(t') \right] \right\} \quad (3.24)$$

$$= \mathbb{I} + \frac{(-i)}{1!} \int_0^t dt_1 \Delta H_I(t_1) + \frac{(-i)^2}{2!} \int_0^t dt_1 \int_0^{t_1} dt_2 T \{ \Delta H_I(t_1) \Delta H_I(t_2) \} + \dots \quad (3.25)$$

where the first line is just a compact shorthand for the second. This expression certainly has the right boundary condition, and it is straightforward to check that it satisfies Eq. (3.23) by explicit differentiation. For $t < 0$, the solution is

$$U_I(t) = T' \left\{ \exp \left[-i \int_0^t dt' \Delta H_I(t') \right] \right\} , \quad (3.26)$$

where T' denotes reverse time ordering

$$T' \{ \phi(x_1) \phi(x_2) \} = \Theta(t_1 - t_2) \phi(x_2) \phi(x_1) + \Theta(t_2 - t_1) \phi(x_1) \phi(x_2) . \quad (3.27)$$

To combine both cases into a simple notation, let us define \tilde{T} to be time ordering for $t > 0$ and reverse time ordering for $t < 0$. Thus,

$$U_I(t) = \tilde{T} \left\{ \exp \left[-i \int_0^t dt' \Delta H_I(t') \right] \right\} . \quad (3.28)$$

In terms of $U(t)$, we now have an analytic (but very complicated) relation between $\phi(t, \vec{x})$ and $\phi_I(t, \vec{x})$.

We can generalize $U(t)$ by defining

$$U_I(t_2, t_1) = U_I(t_2) U_I^\dagger(t_1) = e^{iH_0 t_2} e^{-iH(t_2 - t_1)} e^{-iH_0 t_1} . \quad (3.29)$$

It can be shown that this quantity is equal to

$$U_I(t_2, t_1) = \tilde{T} \left\{ \exp \left[-i \int_{t_1}^{t_2} dt' \Delta H_I(t') \right] \right\} , \quad (3.30)$$

where \tilde{T} is time ordering for $t_2 > t_1$ and reverse time-ordering for $t_2 < t_1$.

The $U_I(t_1, t_2)$ operator has some very nice properties. They include:

1. $U_I(t_1, t_1) = \mathbb{I}$

2. $U_I(t_3, t_2)U_I(t_2, t_1) = U_I(t_3, t_1)$
3. $U_I^\dagger(t_2, t_1) = U_I^{-1}(t_2, t_1) = U_I(t_1, t_2)$

By definition, $U_I(t, 0) = U_I(t)$. This implies trivially that $\phi(x) = U_I^\dagger(t, 0)\phi_I(x)U_I(t, 0)$. Thus, for any product of fields (where $\phi_i = \phi(x_i)$):

$$\begin{aligned} \phi_n \phi_{n-1} \dots \phi_1 &= U_I^\dagger(t_n, 0)\phi_{I_n}U_I(t_n, 0)U_I^\dagger(t_{n-1}, 0)\phi_{I_{n-1}} \dots U_I(t_2, 0)U_I^\dagger(t_1, 0)\phi_{I_1}U_I(t_1, 0) \\ &= U_I^\dagger(t_n, 0)\phi_{I_n}U_I(t_n, t_{n-1})\phi_{I_{n-1}} \dots \phi_{I_2}U_I(t_2, t_1)\phi_{I_1}U_I(t_1, 0) . \end{aligned} \quad (3.31)$$

In this expression, the U_I factors can be thought of as transfer operators that connect the interaction-picture operators at different times.

3.2.2 To Infinity and Beyond

Our goal is to compute the time-ordered operator expectation values $\langle \Omega | T \{ \phi_n, \dots, \phi_1 \} | \Omega \rangle$. Our strategy to do so will be to rewrite them in terms of interaction picture fields sandwiched between the H_0 vacuum $|0\rangle$. This will give us something we know how to deal with, since interaction picture fields act on the $|0\rangle$ state exactly like in the free theory. Evaluating them is then just a matter of commuting a bunch of a and a^\dagger operators through each other.

As a first step, let us expand the H_0 eigenstate $|0\rangle$ in terms of energy eigenstates $\{|N\rangle\}$ of the full Hamiltonian H in a clever way. For this, recall that $H_0|0\rangle = 0$ by construction, and let us write E_Ω as the energy of the true vacuum: $H|\Omega\rangle = E_\Omega|\Omega\rangle$. Multiplying $|0\rangle$ by $U_I^{-1}(-\tau, 0)$, we get

$$U_I^{-1}(-\tau, 0)|0\rangle = e^{-iH\tau}e^{iH_0\tau}|0\rangle \quad (3.32)$$

$$= e^{-iH\tau} \left(|\Omega\rangle\langle\Omega| + \sum_{N>\Omega} |N\rangle\langle N| \right) e^{iH_0\tau}|0\rangle \quad (3.33)$$

$$= e^{-iE_\Omega\tau}\langle\Omega|0\rangle|\Omega\rangle + \sum_{N>\Omega} e^{-iE_N\tau}\langle N|0\rangle|N\rangle , \quad (3.34)$$

where in the second line we have inserted unity in the form of a complete set of H eigenstates. This relation simplifies enormously in the limit $\tau \rightarrow \infty(1 - i\epsilon)$, with $0 < \epsilon \ll 1$. Since $E_N > E_\Omega$, the contributions from all the higher modes become exponentially suppressed in this limit. It follows that

$$|\Omega\rangle = \lim_{\tau \rightarrow \infty(1-i\epsilon)} \frac{e^{iE_\Omega\tau}}{\langle\Omega|0\rangle} U_I^{-1}(-\tau, 0)|0\rangle . \quad (3.35)$$

A similar argument gives

$$\langle\Omega| = \lim_{\tau \rightarrow \infty(1-i\epsilon)} \frac{e^{iE_\Omega\tau}}{\langle 0|\Omega\rangle} \langle 0|U_I(\tau, 0) . \quad (3.36)$$

The second step is to combine these results with the expression of Eq. (3.31). This gives

$$\langle \Omega | \phi_n \dots \phi_1 | \Omega \rangle = \lim_{\tau \rightarrow \infty(1-i\epsilon)} \frac{e^{2iE_\Omega \tau}}{|\langle \Omega | 0 \rangle|^2} \langle 0 | U_I(\tau, 0) U_I^\dagger(t_n, 0) \phi_{I_n} \dots \phi_{I_1} U_I(t_1, 0) U_I^{-1}(-\tau, 0) | 0 \rangle \quad (3.37)$$

Making use of the properties of the $U_I(t, t')$ operators, it is straightforward to show that

$$U_I(\tau, 0) U_I^\dagger(t_n, 0) = U_I(\tau, t_n), \quad U_I(t_1, 0) U_I^{-1}(-\tau, 0) = U_I(t_1, -\tau^*). \quad (3.38)$$

Thus, we have

$$\langle \Omega | \phi_n \dots \phi_1 | \Omega \rangle = \lim_{\tau \rightarrow \infty(1-i\epsilon)} \frac{e^{2iE_\Omega \tau}}{|\langle \Omega | 0 \rangle|^2} \langle 0 | U_I(\tau, t_n) \phi_{I_n} U_I(t_n, t_{n-1}) \dots \phi_{I_1} U_I(t_1, -\tau^*) | 0 \rangle. \quad (3.39)$$

Everything has contracted nicely.

The last two steps are to deal with the funny prefactor and to focus on time-ordered products of fields. For the prefactor, the normalization condition on $|\Omega\rangle$ implies

$$1 = \langle \Omega | \Omega \rangle \quad (3.40)$$

$$= \lim_{\tau \rightarrow \infty(1-i\epsilon)} \frac{e^{2iE_\Omega \tau}}{|\langle \Omega | 0 \rangle|^2} \langle 0 | U_I(\tau, 0) U_I^{-1}(-\tau, 0) | 0 \rangle \quad (3.41)$$

$$= \lim_{\tau \rightarrow \infty(1-i\epsilon)} \frac{e^{2iE_\Omega \tau}}{|\langle \Omega | 0 \rangle|^2} \langle 0 | U_I(\tau, -\tau^*) | 0 \rangle \quad (3.42)$$

$$= \lim_{\tau \rightarrow \infty(1-i\epsilon)} \frac{e^{2iE_\Omega \tau}}{|\langle \Omega | 0 \rangle|^2} \langle 0 | T \left\{ \exp \left[-i \int_{-\tau}^{\tau} dt' \Delta H_I(t') \right] \right\} | 0 \rangle \quad (3.43)$$

In the last line, we have just applied the solution for $U_I(t, t')$ given in Eq. (3.30).

We can also apply our solution for $U(t, t')$ to the operator in Eq. (3.39). If not for the insertions of ϕ_{I_i} in this operator, we could contract up all the $U(t_i, t_{i-1})$ factors to give a single $U(\tau, -\tau^*)$ factor. Unfortunately, ϕ_{I_i} and $\Delta H_I(t)$ typically do not commute and such a simplification is not possible in general. However, we can make this simplification within a time-ordering symbol, since the ordering of operators inside the symbol does not matter.³ Therefore, we can write

$$\langle \Omega | T \{ \phi_n \dots \phi_1 \} | \Omega \rangle = \lim_{\tau \rightarrow \infty(1-i\epsilon)} \frac{\langle 0 | T \left\{ \phi_{I_1} \dots \phi_{I_n} \exp \left[-i \int_{-\tau}^{\tau} dt' \Delta H_I(t') \right] \right\} | 0 \rangle}{\langle 0 | T \left\{ \exp \left[-i \int_{-\tau}^{\tau} dt' \Delta H_I(t') \right] \right\} | 0 \rangle}. \quad (3.44)$$

This result will be used a lot, and we will sometimes call it our *master formula*.

We will show in **notes-04** that time-ordered vacuum matrix elements of fields are precisely what are needed to compute scattering rates. Since they come up so much, they are given a special name and an abbreviated notation. The time-ordered product of n fields is called the n -point function $G^{(n)}(x_1, \dots, x_n)$, and is often written as

$$\langle \Omega | T \{ \phi_1 \dots \phi_n \} | \Omega \rangle = G^{(n)}(x_1, \dots, x_n) = \langle \phi_1 \dots \phi_n \rangle. \quad (3.45)$$

Onward!

³After all, their order will be completely determined by the time ordering, e.g. $T\{\phi_1\phi_2\} = T\{\phi_2\phi_1\}$.

3.2.3 Computing Stuff, Finally

After all this formalism, let's actually compute something. The way the exact result of Eq. (3.44) is usually too complicated to compute exactly. Instead, it is usually computed perturbatively by expanding it in powers of ΔH_I . The right-hand side of Eq. (3.44) is written in terms of interaction-picture fields and the H_0 vacuum $|0\rangle$. This is nice because the interaction picture fields have exactly the same expansion in terms of the a and a^\dagger ladder operators as a free field, Eq. (3.17). Since we know how these act on $|0\rangle$ and how they commute with each other, we know how to compute all the matrix elements.

To be concrete, let us specialize to the case of

$$\Delta H = \int d^3x \frac{g}{3!} \phi^3. \quad (3.46)$$

It follows that $\Delta H_I(t) = \int d^3x \frac{g}{3!} \phi_I^3(t, \vec{x})$. With this in mind, we will evaluate Eq. (3.44) as an expansion in the coupling g .

The easiest thing to compute is the expectation value of the identity operator. For this, Eq. (3.44) simply gives

$$\langle \mathbb{I} \rangle = 1. \quad (3.47)$$

Not so bad at all.

A slightly more challenging quantity is the 2-point function,

$$\langle \phi_1 \phi_2 \rangle = \frac{\langle 0|T \{ \phi_{I_1} \phi_{I_2} \exp [-i \int d^4z \frac{g}{3!} \phi_I^3(z)] \} |0\rangle}{\langle 0|T \{ \exp [-i \int d^4z \frac{g}{3!} \phi_I^3(z)] \} |0\rangle}. \quad (3.48)$$

To leading non-trivial order, which is g^0 in this case, the numerator is

$$\langle 0|T \{ \phi_I(x_1) \phi_I(x_2) \} |0\rangle = D_F(x_1 - x_2), \quad (3.49)$$

while the denominator is equal to unity. There are higher-order contributions, but it is reassuring that we reproduce the free theory result at lowest order.

A quantity that only arises at order g^1 in this theory is the 3-point function. Evaluating it at this order,

$$G^{(3)}(x_1, x_2, x_3) = \langle 0|T \{ \phi_I(x_1) \phi_I(x_2) \phi_I(x_3) \left[1 - i \int d^4z \frac{g}{3!} \phi_I^3(z) \right] \} |0\rangle + \dots \quad (3.50)$$

$$= 0 + \frac{(-i)g}{3!} \int d^4z \langle 0|T \{ \phi_I(x_1) \phi_I(x_2) \phi_I(x_3) \phi_I^3(z) \} |0\rangle + \dots \quad (3.51)$$

The terms omitted in each line are higher-order in g (g^2 or higher). In particular, at this order we can set the denominator to unity. The second line can be evaluated by expanding the fields in terms of ladder operators, arranging their time ordering, and commuting the ladder operators to put everything in normal order. This requires a lot of brute-force calculation to do in general, but fortunately there is a systematic way of computing them called Wick's Theorem.

3.3 Wick's Theorem

To compute the vacuum matrix elements of time-ordered products of interacting fields, Eq. (3.44) tells us we need to manipulate products of a and a^\dagger operators sandwiched between $\langle 0|$ and $|0\rangle$. This is straightforward to do in principle, but the many different terms that arise make it very tedious in practice. *Wick's Theorem* is way to reduce the amount of work.

Wick's Theorem is a statement about free fields that also applies to fields in the interaction picture. In words, it is

$$T \{ \phi(x_1) \dots \phi(x_n) \} = N \{ \phi(x_1) \dots \phi(x_n) + \text{all contractions} \} . \quad (3.52)$$

We will explain what we mean by a contraction below, but for now let us just say that it takes two field operators and turns them into a number. The power of this formula is that a normal-ordered operator vanishes when sandwiched between $\langle 0|$ and $|0\rangle$ unless it is proportional to the identity. This means that only the terms on the right-hand side that are fully contracted, such that no field operators are left, can contribute to the vacuum expectation value.

To describe Wick's theorem, it will be useful to split up the free-field operator into two pieces:

$$\phi(x) = \int \widetilde{dk} \left[a(\vec{k}) e^{-ik \cdot x} + a^\dagger(\vec{k}) e^{ik \cdot x} \right] = \phi_-(x) + \phi_+(x) , \quad (3.53)$$

with⁴

$$\phi_-(x) = \int \widetilde{dk} a(\vec{k}) e^{-ik \cdot x}, \quad \phi_+(x) = \int \widetilde{dk} a^\dagger(\vec{k}) e^{ik \cdot x} . \quad (3.54)$$

From this definition, we find that

$$[\phi_-(x_1), \phi_-(x_2)] = 0 = [\phi_+(x_1), \phi_+(x_2)] , \quad (3.55)$$

together with

$$\phi_-^\dagger(x) = \phi_+(x) , \quad (3.56)$$

as well as

$$[\phi_-(x_1), \phi_+(x_2)] = \int \widetilde{dk} e^{-ik \cdot (x_1 - x_2)} := D(x_1 - x_2) . \quad (3.57)$$

In terms of ϕ_- and ϕ_+ , a product of fields is normal-ordered if and only if all the ϕ_- pieces lie to the right of all the ϕ_+ pieces.

⁴Note that my + and - are flipped relative to P&S.

Since $T\{\mathbb{I}\} = N\{\mathbb{I}\}$ and $T\{\phi(x)\} = N\{\phi(x)\}$ are both trivial, let's look at the product of two fields:

$$T\{\phi(x_1)\phi(x_2)\} = \Theta(t_1 - t_2) [\phi_-(x_1) + \phi_+(x_1)] [\phi_-(x_2) + \phi_+(x_2)] \quad (3.58)$$

$$+ \Theta(t_2 - t_1) [\phi_-(x_2) + \phi_+(x_2)] [\phi_-(x_1) + \phi_+(x_1)]$$

$$= \Theta_{12} (\phi_{1-}\phi_{2-} + \phi_{1+}\phi_{2+} + \phi_{1+}\phi_{2-} + \phi_{2+}\phi_{1-} + [\phi_{1-}, \phi_{2+}]) \quad (3.59)$$

$$+ \Theta_{21} (\phi_{2-}\phi_{1-} + \phi_{2+}\phi_{1+} + \phi_{2+}\phi_{1-} + \phi_{1+}\phi_{2-} + [\phi_{2-}, \phi_{1+}])$$

$$= \phi_{1-}\phi_{2-} + \phi_{1+}\phi_{2-} + \phi_{2+}\phi_{1-} + \phi_{1+}\phi_{2+} \quad (3.60)$$

$$+ \Theta(t_1 - t_2)D(x_1 - x_2) + \Theta(t_2 - t_1)D(x_2 - x_1)$$

$$= N\{\phi_1\phi_2\} + D_F(x_1 - x_2) . \quad (3.61)$$

In the third line, we have used $\Theta_{12} + \Theta_{21} = 1$, while in the fourth we have combined the two D functions into a D_F .⁵ Let us now define the contraction of two fields to be

$$\overline{\phi(x_1)\phi(x_2)} = D_F(x_1 - x_2) . \quad (3.62)$$

With this in place, we have

$$T\{\phi(x_1)\phi(x_2)\} = N\{\phi(x_1)\phi(x_2) + \overline{\phi(x_1)\phi(x_2)}\} . \quad (3.63)$$

Note that since D_F is just a function proportional to the unit operator, we can freely move it inside the normal ordering. This is consistent with our statement of Wick's Theorem.

Doing the same calculation for three fields, one obtains

$$T\{\phi_1\phi_2\phi_3\} = N\{\phi_1\phi_2\phi_3 + \phi_1\overline{\phi_2\phi_3} + \phi_2\overline{\phi_3\phi_1} + \phi_3\overline{\phi_1\phi_2}\} \quad (3.64)$$

$$= N\{\phi_1\phi_2\phi_3 + \phi_1D_F(x_2 - x_3) + \phi_2D_F(x_1 - x_3) + \phi_3D_F(x_1 - x_2)\} . \quad (3.65)$$

For four fields,

$$\begin{aligned} T\{\phi_1\phi_2\phi_3\phi_4\} &= N\{\phi_1\phi_2\phi_3\phi_4 + \overline{\phi_1\phi_2}\phi_3\phi_4 + \overline{\phi_1\phi_3}\phi_2\phi_4 + \overline{\phi_1\phi_4}\phi_2\phi_3 \\ &\quad + \overline{\phi_2\phi_3}\phi_1\phi_4 + \overline{\phi_2\phi_4}\phi_1\phi_3 + \overline{\phi_3\phi_4}\phi_1\phi_2 \\ &\quad + \overline{\phi_1\phi_2}\overline{\phi_3\phi_4} + \overline{\phi_1\phi_3}\overline{\phi_2\phi_4} + \overline{\phi_1\phi_4}\overline{\phi_2\phi_3}\} . \end{aligned} \quad (3.66)$$

Oof.

The proof of Wicks' theorem goes by induction. Suppose it holds for any product of n fields. For $(n + 1)$ fields labelled such that $t_1 > t_2 > \dots$, we get

$$T\{\phi_1\phi_2 \dots \phi_{n+1}\} = \phi_1 T\{\phi_2 \dots \phi_{n+1}\} \quad (3.67)$$

$$= (\phi_{1+} + \phi_{1-}) N\{\phi_2 \dots \phi_{n+1} + \text{contractions}'\} \quad (3.68)$$

where $\text{contractions}'$ means all contractions that do not involve ϕ_1 . In this expression, the ϕ_{1+} term is already in normal order, so we only need to move ϕ_{1-} through to the right. This will

⁵Check this for yourself by expanding out the fields in the definition of D_F .

eventually produce something that is normal-ordered together with a bunch of contractions, noting that $[\phi_-(x_1), \phi_+(x_i)] = D_F(x_1 - x_i)$ since we have $t_1 > t_i$ for all i . All that remains to show is that all possible contractions involving ϕ_1 are produced. We can do this by giving an algorithm to build a given contraction given ϕ_{1-} on the left and the terms already assumed to be present in the inductive step. I won't do this explicitly, but it isn't too difficult with a bit of fiddling.

Our statement and (partial) proof of Wick's Theorem was for free scalar fields. However, it also applies to scalar fields in the interaction picture since they obey the same commutation relations and have the same ladder operator expansions as free fields. Therefore Wick's theorem allows us to systematically evaluate the operators that arise in the vacuum expectation values of time-ordered products of fields.

3.4 A First Look at Feynman Diagrams and Rules

The contractions that arise in computing n -point functions can be represented by diagrams. These *Feynman diagrams* help to keep track of all the possible terms. Even better, it is possible to assign a value to each Feynman diagram. Computing n -point functions can thus be reduced to drawing a set of pictures and computing their values using a set of *Feynman rules*.

To illustrate how Feynman diagrams work, let us return to the interacting scalar theory with $\Delta V = g\phi^3/3!$ and evaluate the leading-order expression for the 3-point function we found in Eqs. (3.50,3.51). Applying Wick's Theorem,

$$\begin{aligned} \langle 0|T\{\phi_I(x_1)\phi_I(x_2)\phi_I(x_3)\phi_I^3(z)\}|0\rangle = & \quad (3.69) \\ (3!)D_F(x_1 - z)D_F(x_2 - z)D_F(x_3 - z) + 3D_F(x_1 - z)D_F(x_2 - x_3)D_F(z - z) \\ + 3D_F(x_2 - z)D_F(x_1 - x_3)D_F(z - z) + 3D_F(x_3 - z)D_F(x_1 - x_2)D_F(z - z) \end{aligned}$$

Note that only fully contracted terms in the Wick expansion contribute to the matrix element. The numerical prefactors are just the number of ways to get each specific contraction.

We can represent this set of contractions in pictures. To each term, we associate a distinct Feynman diagram. The set of diagrams corresponding to Eq. (3.69) is shown in Fig. 3.1. The rules for drawing a diagram are as follows:

1. For each coordinate x_i , draw a dot.
2. A contraction producing a factor of $D_F(x_i - x_j)$ is represented by a line connecting the dot for x_j to the dot for x_i .
3. Each factor of $D_F(z - z)$ is depicted as a loop connecting z to itself.

The utility of these rules is that we can invert them, using diagrams to figure out all the possible contractions, and then associating a number to each diagram. Instead of wading through a mire of raising and lowering operators, we only need to connect up some dots.

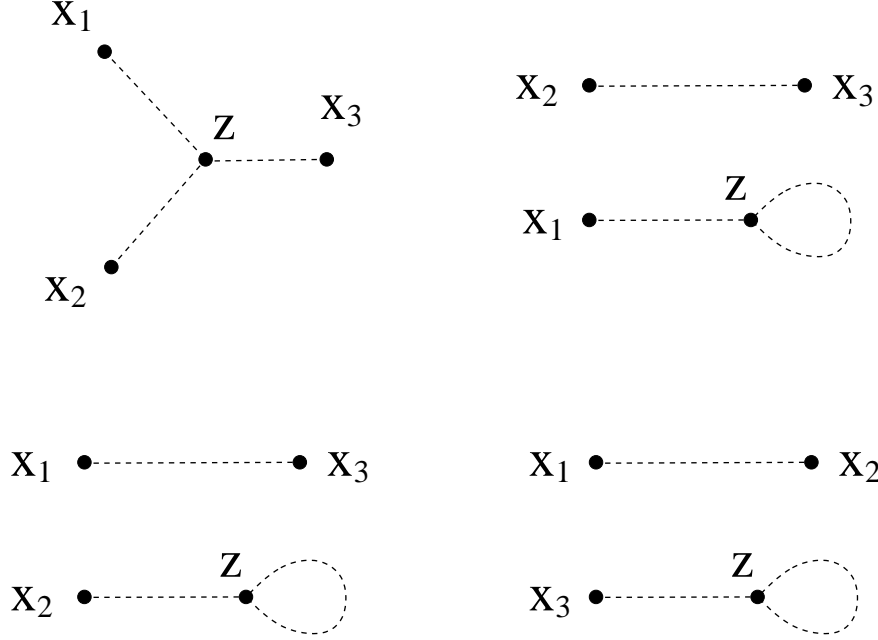


Figure 3.1: Feynman diagrams for the 3-point function at $\mathcal{O}(g)$ of Eq. (3.69).

Let's now build up a set of *Feynman rules* for computing time-ordered vacuum matrix elements of fields in the $g\phi^3/3!$ theory. These rules will specify how to draw a set of Feynman diagrams for a given matrix element and assign a numerical value to each diagram. They are based on the master formula of Eq. (3.44), which we treat by expanding the exponentials it contains in a power series in g up to some fixed order g^M . At this order, the perturbation theory estimate for the matrix element is given by the sum of each of the individual g^0, g^1, \dots, g^M contributions. If g is small enough, keeping only the first few terms in the expansion should give a good approximation to the truth.

The Feynman rules to compute the g^M contribution to the n -point function $G^{(n)}(x_1, \dots, x_n) = \langle \Omega | T \{ \phi(x_1) \dots \phi(x_n) \} | \Omega \rangle$ in the $g\phi^3/3!$ theory are:

1. Draw a dot for each x_i coordinate. We call these *external points*.
2. Draw another M dots and associate a coordinate z_j ($j = 1, 2, \dots, M$) to each of them. We call these *vertices*.
3. Draw Feynman diagrams by connecting the dots with lines in all possible ways following two simple rules:
 - a) Each external point has a single line connected to it.
 - b) Each vertex has three lines connected to it.
4. Remove all the diagrams in which there is a *vacuum bubble* – a vertex point that not connected by any path to any of the external points.
5. Write the number corresponding to each Feynman diagram:
 - a) For every vertex with coordinate z_j , write a factor of $-ig \int d^4 z_j$.

- b) For every line connecting points a and b (external or vertex), add a factor of the Feynman propagator $D_F(a - b)$.
- c) Any diagram with an unconnected dot or an unpaired line is equal to zero.
- d) Multiply the result for each diagram by $1/M!$ and a *symmetry factor*.

The net g^M contribution is then the sum of all the diagrams.

To see how this works in practice, let us begin with the 2-point function and work to second order (g^2) in the perturbative expansion. From our master formula, we have

$$G^{(2)}(x_1, x_2) = \langle 0|T \left\{ \phi_1 \phi_2 \left[1 + \frac{(-i)g}{3!} \int d^4z \phi_z^3 + \frac{(-i)^2 g^2}{3!3!2!} \int d^4z_1 \int d^4z_2 \phi_{z_1}^3 \phi_{z_2}^3 \right] \right\} |0\rangle \quad (3.70)$$

$$\left/ \langle 0|T \left\{ 1 + \frac{(-i)g}{3!} \int d^4z \phi_z^3 + \frac{(-i)^2 g^2}{3!3!2!} \int d^4z_1 \int d^4z_2 \phi_{z_1}^3 \phi_{z_2}^3 \right\} |0\rangle \right.$$

Following the rules and drawing the Feynman diagrams, we find the set shown in Fig. 3.2. At leading order g^0 , we only have the first diagram in the figure and associate to it the number

$$G^{(2)}(x_1, x_2)|_{g^0} = D_F(x_1 - x_2) . \quad (3.71)$$

This matches what we would have found from applying Wick's Theorem to the matrix element. There are no contributions proportional to g^1 because it is impossible to connect up all the dots and all the lines. Again, this matches what the matrix element would produce. Going to g^2 , we find the second row of diagrams in Fig. 3.2. Their sum is

$$G^{(2)}(x_1, x_2)|_{g^2} = \quad (3.72)$$

$$\left[\frac{1}{2} \right] \left(\frac{2!}{2!} \right) (-ig)^2 \int d^4z_1 \int d^4z_2 D_F(x_1 - z_1) D_F(z_1 - z_2) D_F(z_1 - z_2) D_F(z_2 - x_2)$$

$$+ \left[\frac{1}{2} \right] \left(\frac{2!}{2!} \right) (-ig)^2 \int d^4z_1 \int d^4z_2 D_F(x_1 - z_1) D_F(z_1 - z_2) D_F(z_2 - z_2) D_F(z_1 - x_2)$$

$$+ \left[\frac{1}{4} \right] \left(\frac{2!}{2!} \right) (-ig)^2 \int d^4z_1 D_F(x_1 - z_1) D_F(z_1 - z_1) \int d^4z_2 D_F(x_2 - z_2) D_F(z_2 - z_2)$$

The factors of $(2!/2!)$ come from the $(1/2!)$ in the expansion of the exponential together with the fact that since z_1 and z_2 are integrated over, diagrams differing only by the permutation of z_1 and z_2 have the same value. The numerical prefactors in the square brackets $(1/2, 1/2, 1/4)$ are the symmetry factors. Note as well that we have omitted the lower row of diagrams in Fig. 3.2 (surrounded by dotted boxes) because they have a vacuum bubble, in which one or more vertices are not connected to any of the external points in any way. The full 2-point function in the interacting theory up to order g^2 in perturbation theory is therefore

$$G^{(2)}(x_1, x_2) = G^{(2)}(x_1, x_2)|_{g^0} + G^{(2)}(x_1, x_2)|_{g^2} + \dots . \quad (3.73)$$

This is just a number to be evaluated.

Most of the Feynman rules are easy to relate to what's going on in matrix element after applying Wick's Theorem. However, rules 4 and 5d) might seem a bit strange. For rule 4,

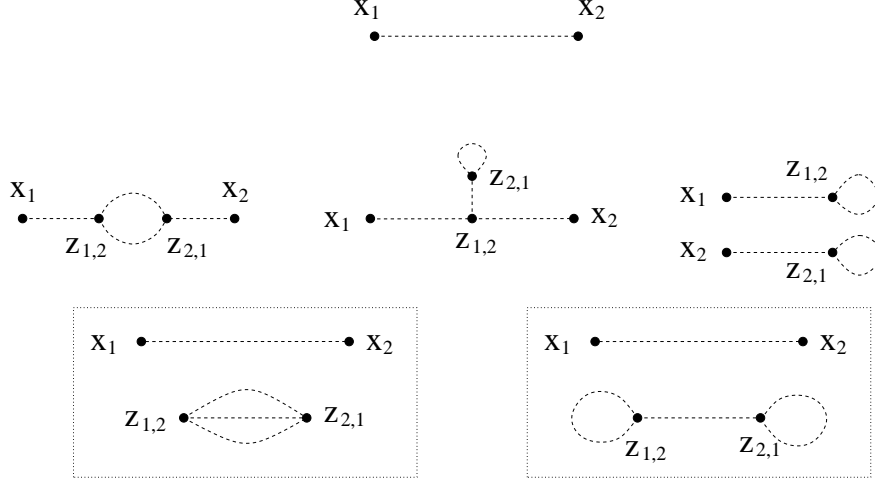


Figure 3.2: Feynman diagrams for the 2-point function at $\mathcal{O}(g^2)$.

the reason we can drop these diagrams is that they are cancelled off by the expansion of the denominator in Eq. (3.44). We can see this explicitly here by studying the matrix element. Applying Wick's theorem to the denominator in Eq. (3.70), we get

$$\begin{aligned}
 \text{Denom.} &= 1 + \left[\frac{1}{12} \right] (-i)^2 g^2 \int d^4 z_1 \int d^4 z_2 [D_F(z_1 - z_2)]^3 \\
 &\quad + \left[\frac{1}{8} \right] (-i)^2 g^2 \int d^4 z_1 \int d^4 z_2 D_F(z_1 - z_2) D_F(z_1 - z_1) D_F(z_2 - z_2) \\
 &:= (1 + g^2 \Delta) .
 \end{aligned} \tag{3.74}$$

Using $1/(1 + g^2 \Delta) = 1 - g^2 \Delta + g^4 \Delta^2 + \dots$, the net result up to order g^2 is

$$G^{(2)}(x_1, x_2) = [D_F(x_1 - x_2) + g^2(\dots)] \times (1 - g^2 \Delta) \tag{3.75}$$

$$= D_F(x_1 - x_2)(1 - g^2 \Delta) + g^2(\dots) , \tag{3.76}$$

where the $g^2(\dots)$ term corresponds to all the stuff in the numerator of the matrix element proportional to g^2 . Some of the terms in the numerator will be the diagrams in the bottom row of Fig. 3.2 that contain vacuum bubbles. Using the Feynman rules, it is easy to show that these diagrams with vacuum bubbles sum to $D_F(x_1 - x_2)\Delta$. These diagrams are therefore cancelled off at this order by the expansion of the denominator. It turns out that such a cancellation between the denominator and diagrams in the numerator containing vacuum bubbles is a general feature, and a proof can be found in Peskin & Schroeder [1]. This explains the origin of rule 4.

Some diagrams also require an additional symmetry factor. In many cases, summing over all the contractions cancels off the $(1/3!)$ in our definition of $\Delta V = g\phi^3/3!$. However, sometimes this cancellation is incomplete and we need to correct for it with an additional factor, called the symmetry factor of the diagram. In general, it is equal to one divided by the number of ways that internal lines can be reconnected to give the same diagram [1, 2].

As a practical matter, most people just work out the number of contractions corresponding to each diagram by counting contractions using Wick's theorem.

Finally, let us also mention that the denominator factor of $2!$ from the expansion of the exponential cancelled against the factor of $2!$ arising from the fact that for each diagram with fixed z_1 and z_2 , there was a corresponding diagram with $z_1 \leftrightarrow z_2$ having the same numerical value. This is also a general feature at any order because all the internal z_i coordinates are integrated over. Thus, we can slightly modify our Feynman rule 5d) and leave out the $1/M!$ factor (at order g^M) with the understanding that we are only to include those diagrams that remain distinct when the z_i coordinates are permuted.

Chapter 4

Feynman Rules and Scattering

We now have a perturbative prescription to compute the vacuum matrix elements of time-ordered products of fields. While it's nice to be able to compute something, even approximately, what we would really like is to do is predict physically observable quantities. In this note we will show how to relate time-ordered matrix elements to quantum amplitudes for particle scattering.

4.1 Asymptotic States

Our first goal is to connect n -point functions to incoming and outgoing particles travelling from and to spacetime infinity. For this, we will use two primary results: the Källén-Symanzik spectral decomposition, and the Lehmann-Symanzik-Zimmermann formula.

4.1.1 Spectral Decomposition

To begin, let us think about the effect of translations and Lorentz transformations on the scalar fields in the interacting theory (with some unspecified $\Delta V(\phi)$). Just like in the free theory, spacetime translations are symmetries of the theory and there are corresponding conserved 4-momentum operators $P^\mu = (H, \vec{P})$. In particular, we have

$$\phi(x) = e^{iP \cdot x} \phi(0) e^{-iP \cdot x} . \quad (4.1)$$

Since spacetime translations commute and the operators are Hermitian, there exists a basis of P^μ eigenstates that spans the Hilbert space.

We have already made a couple of assumptions about the structure of these states. The first is that there is a vacuum state $|\Omega\rangle$ such that¹

$$P^\mu |\Omega\rangle = 0 . \quad (4.2)$$

¹Compared to before, we have also set $E_\Omega = 0$. This can be arranged by a nice choice for Λ in ΔV . Furthermore, having $E_\Omega \neq 0$ doesn't change any of our arguments.

The second assumption is that the next set of states up in energy are isolated one-particle states $|\vec{p}\rangle$ with

$$P^\mu |\vec{p}\rangle = p^\mu |\vec{p}\rangle , \quad (4.3)$$

with $p^0 = \sqrt{\vec{p}^2 + M^2}$ for some constant $M^2 > 0$. More precisely, we assume there is a unique one-particle state with this value of M for every possible value of the 3-momentum \vec{p} . We will normalize these states in the same way as the momentum states of the free theory,

$$\langle \vec{p} | \vec{k} \rangle = (2\pi)^3 2p^0 \delta^{(3)}(\vec{p} - \vec{k}) . \quad (4.4)$$

For the remaining 4-momentum eigenstates $|\psi'_{\vec{p}}\rangle$ above the one-particle states, we make no further assumptions aside from a mass gap. By this, we mean that

$$P^\mu |\psi'_{\vec{p}}\rangle = p^\mu_{\psi'} |\psi'_{\vec{p}}\rangle \quad (4.5)$$

with $p^0_{\psi'} = E_{\psi'}(\vec{p}) = \sqrt{\vec{p}^2 + M_{\psi'}^2}$ with $M_{\psi'}^2 > M^2$. Note that since $P^2 = E^2 - \vec{p}^2$ is Lorentz-invariant, it must be equal to a constant and thus the energy must take this general form. Although we don't specify what the $|\psi'_{\vec{p}}\rangle$ states are, they include multiparticle states with total 3-momentum $\vec{p}_{\psi'}$ and possibly also single-particle bound states with mass greater than M . Collectively, we will refer to all the states above the vacuum (one-particle and higher) as $\{|\psi_{\vec{p}}\rangle\}$

It is also useful to think about the effects of Lorentz transformations (boosts and rotations) on the system. These act on spacetime according to

$$x^\mu \rightarrow x'^\mu = \Lambda^\mu_\nu x^\nu . \quad (4.6)$$

We will study these in more detail later on in the course, but for now all we need to know is that such transformations can be implemented by a unitary operator $U(\Lambda)$ on the Hilbert space. This operator does not alter the vacuum (by assumption):

$$U|\Omega\rangle = |\Omega\rangle . \quad (4.7)$$

On the field operator, we have

$$U(\Lambda)\phi(x)U^{-1}(\Lambda) = \phi(\Lambda^{-1}x) . \quad (4.8)$$

In particular, this implies $U(\Lambda)\phi(0)U^{-1}(\Lambda) = \phi(0)$. For any state carrying momentum (one-particle or otherwise), we can always rewrite it as a boost operator $U(\Lambda_{\vec{p}})$ acting on a state with zero 3-momentum:

$$|\psi_{\vec{p}}\rangle = U(\Lambda_{\vec{p}})|\psi_{\vec{0}}\rangle . \quad (4.9)$$

In this way, we can think of all the different one-particle states as boosts of the state with a single particle at rest, $p^\mu = (M, \vec{0})$.

The completeness of the set of momentum eigenstates can be written as

$$\mathbb{I} = |\Omega\rangle\langle\Omega| + \sum_{\psi} \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\psi}(\vec{p})} |\psi_{\vec{p}}\rangle\langle\psi_{\vec{p}}|. \quad (4.10)$$

Here, the sum over ψ includes the one-particle and other states, and it only runs over states that are not related to each other by a Lorentz transformation. In other words, it runs only over the $|\psi_{\vec{0}}\rangle$ states since the rest can be obtained by boosting.

Let us now insert this resolution of the identity into the matrix element of a pair of fields:

$$\langle\Omega|\phi(x_1)\phi(x_2)|\Omega\rangle = \langle\Omega|\phi(0)e^{-iP\cdot(x_1-x_2)}\phi(0)|\Omega\rangle \quad (4.11)$$

$$= |\langle\Omega|\phi(0)|\Omega\rangle|^2 + \sum_{\psi} \int \frac{d^3p}{(2\pi)^3} \frac{e^{-ip_{\psi}\cdot(x_1-x_2)}}{2E_{\psi}(\vec{p})} |\langle\Omega|\phi(0)|\psi_{\vec{p}}\rangle|^2 \quad (4.12)$$

$$= 0 + \sum_{\psi} \int \frac{d^3p}{(2\pi)^3} \frac{e^{-ip_{\psi}\cdot(x_1-x_2)}}{2E_{\psi}(\vec{p})} |\langle\Omega|\phi(0)|\psi_0\rangle|^2 \quad (4.13)$$

In the last line we have used $|\psi_{\vec{p}}\rangle = U(\Lambda_{\vec{p}})|\psi_0\rangle$ and we have applied our assumption that $\langle\Omega|\phi(0)|\Omega\rangle = 0$.

Using $E_{\psi}(\vec{p}) = \sqrt{\vec{p}^2 + M_{\psi}^2}$ and putting in a time ordering, Eq. (4.13) implies

$$\langle\Omega|T\{\phi_1\phi_2\}|\Omega\rangle = \sum_{\psi} \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - M_{\psi}^2 + i\epsilon} e^{-ip\cdot(x_1-x_2)} |\langle\Omega|\phi(0)|\psi_0\rangle|^2 \quad (4.14)$$

$$= \int_0^{\infty} \frac{ds}{2\pi} \rho(s) D_F(x_1 - x_2; s), \quad (4.15)$$

where $D_F(x_1 - x_2; s)$ is the Feynman propagator for a free field of mass $m^2 = s$, and $\rho(s)$ is defined to be

$$\rho(s) = \sum_{\psi} (2\pi)\delta(s - M_{\psi}^2) |\langle\Omega|\phi(0)|\psi_0\rangle|^2 \quad (4.16)$$

The result of Eq. (4.15) is called the Källén-Lehmann spectral representation. In the free theory, we would just have $\rho(s) = 2\pi\delta(s - m^2)$ and the 2-point function reduces to the Feynman propagator with $s = m^2$. In the interacting theory, $\rho(s)$ is a non-trivial *spectral function* that characterizes the excitations of the theory. With our assumption of an isolated one-particle state of mass M , it takes the form

$$\rho(s) = 2\pi Z \delta(s - M^2) + \bar{\rho}(s), \quad (4.17)$$

where $Z = |\langle\Omega|\phi(0)|\vec{p}\rangle|^2 > 0$. We illustrate the form of $\rho(s)$ under this assumption in Fig. 4.1. Above the one-particle state, there could be isolated bound states, and there will definitely be a continuum of multi-particle states at $s \geq (2M)^2$. As a function in the complex s plane, distinct particles correspond to isolated poles, while the continuum of multiple particles corresponds to a branch cut along the real line.

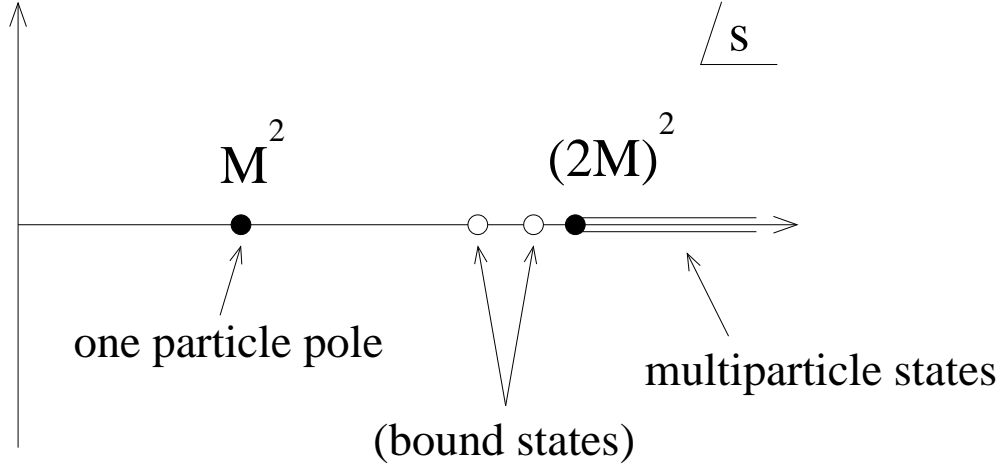


Figure 4.1: Schematic depiction of the spectral function in a theory with a weak interaction.

4.1.2 Lehmann, Symanzik, and Zimmermann (LSZ)

The LSZ reduction formula is one of the key tools in perturbative quantum field theory. It relates the vacuum matrix elements of time-ordered products of field operators to the matrix elements for particle scattering. We will not go through the proof of the formula (which can be found in Peskin & Schroeder [1]), but we will try to motivate it.

Consider first the spectral representation of the 2-point function. Let us Fourier transform the result to momentum space:

$$\int d^4x e^{ik \cdot x} \langle \Omega | T \{ \phi(x) \phi(0) \} | \Omega \rangle = \int_0^\infty \frac{ds}{2\pi} \rho(s) \frac{i}{k^2 - s + i\epsilon}. \quad (4.18)$$

If we now apply our assumption of an isolated one-particle state, this becomes

$$\int d^4x e^{ik \cdot x} \langle \Omega | T \{ \phi(x) \phi(0) \} | \Omega \rangle = \frac{iZ}{k^2 - M^2 + i\epsilon} + \int_{>M^2}^\infty \frac{ds}{2\pi} \bar{\rho}(s) \frac{i}{k^2 - s + i\epsilon}. \quad (4.19)$$

Viewed as a function in the complex k^0 plane, this quantity has an isolated pole at $k^0 = +\sqrt{\vec{k}^2 + M^2} - i\epsilon$, corresponding to the physical energy of a one-particle state with momentum \vec{k} and mass M , as well as a more complicated and undertermined structure further out in the k^0 plane. This means that if we want to focus on the one-particle state and ignore the rest of the junk that comes up, we just need to isolate the k^0 pole.

The LSZ reduction formula is just a generalization of this observation. It states that the *connected amplitude* for an initial state with m well-separated particles in the initial state at $t \rightarrow -\infty$ with 3-momenta $\vec{k}_1, \vec{k}_2, \dots, \vec{k}_m$ to go to a final state with n well-separated particles

at $t \rightarrow +\infty$ with 3-momenta $\vec{p}_1, \vec{p}_2, \dots, \vec{p}_n$ is given by

$$\langle \vec{p}_1 \dots \vec{p}_n | \vec{k}_1 \dots \vec{k}_m \rangle_c = \tag{4.20}$$

$$\begin{aligned} & \left(\lim_{k_1^2 \rightarrow M^2} \dots \lim_{k_m^2 \rightarrow M^2} \right) \left(\lim_{p_1^2 \rightarrow M^2} \dots \lim_{p_n^2 \rightarrow M^2} \right) \frac{i^{n+m}}{(\sqrt{Z})^{(m+n)}} \\ & \int d^4 z_1 e^{-ik_1 \cdot z_1} (\partial_{z_1}^2 + M^2) \dots \int d^4 z_m e^{-ik_m \cdot z_m} (\partial_{z_m}^2 + M^2) \\ & \int d^4 x_1 e^{ip_1 \cdot x_1} (\partial_{x_1}^2 + M^2) \dots \int d^4 x_n e^{ip_n \cdot x_n} (\partial_{x_n}^2 + M^2) \\ & \langle \Omega | T \{ \phi(x_1) \dots \phi(x_n) \phi(z_1) \dots \phi(z_m) \} | \Omega \rangle \\ & = \left(\lim_{k_1^2 \rightarrow M^2} \dots \lim_{k_m^2 \rightarrow M^2} \right) \left(\lim_{p_1^2 \rightarrow M^2} \dots \lim_{p_n^2 \rightarrow M^2} \right) \frac{i^{n+m}}{(\sqrt{Z})^{(m+n)}} \\ & \prod_{i=1}^m \left[\int d^4 z_i e^{-ik_i \cdot z_i} (-k_i^2 + M^2) \right] \prod_{j=1}^n \left[\int d^4 x_j e^{ip_j \cdot x_j} (-p_j^2 + M^2) \right] \\ & \langle \Omega | T \{ \phi(x_1) \dots \phi(x_m) \phi(z_1) \dots \phi(z_n) \} | \Omega \rangle \end{aligned} \tag{4.21}$$

To get the second line, we have integrated by parts in each of the integrals.

As an operational tool, the LSZ formula can be applied to find the matrix element for $\langle \vec{p}_1 \dots \vec{p}_n | \vec{k}_1 \dots \vec{k}_m \rangle_c$ by following a few simple steps. First, compute the $(m+n)$ -point function and take its Fourier transform using $\int d^4 z e^{-ik \cdot z}$ for all incoming particles and $\int d^4 x e^{ip \cdot x}$ for all outgoing ones. Next, identify all the poles at $k_i^2 = M^2$ and $p_j^2 = M^2$. Cancel off the poles in the terms that contain a product of all of them and set everything else to zero. Finally, take the limit $k_i^2 \rightarrow M^2$ and $p_j^2 \rightarrow M^2$. This procedure isolates the part of the $(m+n)$ -point function that corresponds to m isolated initial particles and n isolated final particles. The portion of the $(m+n)$ -point function that does not have all the necessary poles vanishes when it is multiplied by the $(p^2 - M^2)$ factors and the $p^2 \rightarrow M^2$ limit is taken.²

Let us apply this result to the 3-point function in the interacting theory with $\Delta V = g\phi^3/3!$ that we computed to order g^1 in note-3 (see Fig.1). Recall that we had

$$\begin{aligned} G^{(3)}(x_1, x_2, x_3) &= (-ig) \int d^4 z D_F(x_1 - z) D_F(x_2 - z) D_F(x_3 - z) \\ &+ \frac{(-ig)}{2} \int d^4 z D_F(z - z) D_F(x_3 - z) D_F(x_1 - x_2) + (\text{permutations}) . \end{aligned} \tag{4.22}$$

Fourier transforming the first term (with three initial states) produces

$$T_1 = (2\pi)^4 \delta^{(4)}(p_1 + p_2 + p_3) (-ig) \frac{i}{p_1^2 - M^2} \frac{i}{p_2^2 - M^2} \frac{i}{p_3^2 - M^2} . \tag{4.23}$$

This has three poles right where we expect them. On the other hand, the Fourier transform

² I'm being sloppy with the $i\epsilon$ factors here, but feel free to put them in as needed.

of the second term is

$$T_2 = (2\pi)^4 \delta^{(4)}(p_1 + p_2 + p_3) \left[\int \frac{d^4 k}{(2\pi)^4} \frac{i}{k^2 - M^2} \right] \frac{i}{p_1^2 - M^2} \frac{i}{p_3^2 - M^2} (2\pi)^4 \delta^{(4)}(p_3) , \quad (4.24)$$

which disappears when it is multiplied by $(p_1^2 - M^2)(p_2^2 - M^2)(p_3^2 - M^2)$, as do the permutations. Applying the LSZ formula therefore produces

$$\langle \Omega | \vec{p}_1 \vec{p}_2 \vec{p}_3 \rangle_c = (-ig) \times (2\pi)^4 \delta^{(4)}(p_1 + p_2 + p_3) . \quad (4.25)$$

Note that we would have obtained the nearly same result if we had put some of the particles in the final state. The only difference would have been $p_i \rightarrow -p_i$ in the delta function for every final-state particle. The delta function therefore enforces the overall conservation of energy and momentum. It turns out that this is a universal feature.

The terms that vanish when the LSZ formula is applied to the 3-point function (at order g^1) correspond to *disconnected diagrams*. These diagrams contain one or more particles that just pass through without interacting with the others at all. They are not what we are interested in when we compute scattering, and it is a useful feature of the LSZ formula that they are automatically removed. This is also why we said that the resulting matrix element is connected.

4.1.3 The S -Matrix

Our result for matrix elements between incoming and outgoing particles that are well-separated at spacetime infinity is frequently written in terms of an S -matrix. If we restrict ourselves to well-separated sets of particles at $t \rightarrow \pm\infty$, we can map these states in an obvious way to the set of momentum eigenstates of a free field theory. In this picture, the free states at $t \rightarrow -\infty$ are called IN states and those at $t \rightarrow +\infty$ are called OUT states.

The matrix elements of interest for scattering are therefore

$${}_{OUT} \langle \vec{p}_1 \dots \vec{p}_n | \vec{k}_1 \dots \vec{k}_m \rangle_{IN} , \quad (4.26)$$

where the value of the inner product is defined to be equal to the value obtained between the corresponding states in the interacting theory. The S matrix is defined to be the mapping between these two different but equivalent Hilbert spaces:

$$|\vec{k}_1 \dots \vec{k}_m \rangle_{IN} = S |\vec{k}_1 \dots \vec{k}_m \rangle_{OUT} . \quad (4.27)$$

In other words, the S -matrix maps each OUT state to the corresponding element of the IN Hilbert space, as determined by the dynamics of the interacting theory. It is a map from one Hilbert space defined at $t \rightarrow -\infty$ to an equivalent one at $t \rightarrow +\infty$, and it is invertible. In the absence of scattering, the S matrix is unity, which is what we found for the free theory.

The key property of the S matrix is that it is unitary. Note that we have

$$\delta(\{\vec{k}_i\} - \{\vec{p}_i\}) = {}_{IN} \langle \{\vec{k}_i\} | \{\vec{p}_i\} \rangle_{IN} \quad (4.28)$$

$$= {}_{OUT} \langle \{\vec{k}_i\} | S^\dagger S | \{\vec{p}_i\} \rangle_{OUT} . \quad (4.29)$$

The orthogonality of the OUT states implies that $S^\dagger S = \mathbb{I}$. Physically, the unitarity of the S matrix corresponds to the conservation of probability, in the sense that the sum over all the IN to OUT squared matrix elements adds up to one. This is also expected from our assumption that the time evolution in quantum mechanics is unitary.

We will frequently write

$$S = e^{iT} \simeq \mathbb{I} + iT, \quad (4.30)$$

where the T matrix is Hermitian. The unit term in the expansion of the S matrix corresponds to the case of no scattering. We will therefore be interested primarily in the T matrix.

4.2 Feynman Rules in Momentum Space

In the LSZ formula, all the position variables of the n -point function are Fourier transformed. For this reason, it is useful to formulate Feynman rules directly in momentum space, where the position variables do not appear at all. Once we have the transformed n -point function, it is trivial to apply the LSZ formula.

4.2.1 Feynman Rules for the Transformed n -Point Function

We will continue to study the real scalar theory with $\Delta V = g\phi^3/3!$. Let us define the Fourier transform of the n -point function by

$$(2\pi)^4 \delta^{(4)} \left(\sum_{i=1}^n p_i \right) \tilde{G}^{(n)}(p_1, \dots, p_n) = \left(\prod_{i=1}^n \int d^4 x_i e^{-ip_i \cdot x_i} \right) \langle \Omega | T \{ \phi(x_1) \dots \phi(x_n) \} | \Omega \rangle \quad (4.31)$$

The Feynman rules for the order- g^M portion of the momentum space n -point function are:

1. Draw an *external line* for each momentum p_i with one fixed end and one free end.
2. Put in M vertices, each with three lines (with free ends) coming out of it.
3. Assemble all the Feynman diagrams by connecting the free ends of the external lines and the vertex lines to each other in pairs in all possible ways.
4. Remove all the diagrams with vacuum bubbles and any diagram that has one or more lines with an unconnected free end.
5. Assign a value to each diagram:
 - a) Each line gets a propagator factor of $i/(p^2 - M^2 + i\epsilon)$. The 4-momentum p is equal to p_i on external lines. The momentum on any other *internal* line is underdetermined at this point, so call it whatever you like, q_j say.
 - b) Write a factor of $-ig$ for each vertex.

- c) Add a factor of $(2\pi)^4 \delta^{(4)}(\sum k_i)$ for each vertex, where the sum runs over all momenta (internal or external) flowing into the vertex, and $k_i \rightarrow -k_i$ if the momentum is flowing out of the vertex. Also, if a pair of external lines p_i and p_j are connected, add a factor of $(2\pi)^4 \delta^{(4)}(p_i + p_j)$. The delta functions arise automatically from the Fourier transforms, and they have the effect of imposing the conservation of 4-momentum at each vertex.
- d) Integrate over each of the internal momenta: $\int d^4 q_j / (2\pi)^4$.
- e) Multiply each diagram by its symmetry factor.

The resulting sum of the diagrams is the order- g^M contribution to $\tilde{G}^{(n)}(p_1, \dots, p_n)$ times the overall delta function.

All this may sound complicated, but it is really pretty easy once you see a few examples and get the hang of it. Consider first the 2-point function at leading order:

$$(2\pi)^4 \delta^{(4)}(p_1 + p_2) \tilde{G}^{(2)}(p_1, p_2) = (2\pi)^4 \delta^{(4)}(p_1 + p_2) \frac{i}{p_1^2 - M^2} \quad (4.32)$$

$$\Rightarrow \tilde{G}^{(2)}(p) = \frac{i}{p^2 - M^2} . \quad (4.33)$$

Note that this would normally depend on two arguments, but since $p_1 + p_2 = 0$ is enforced by the overall delta function, it is conventional to write it with just a single argument $p = p_1 = -p_2$. Going to order g^2 , the relevant diagrams are just like in Fig.2 of notes-03. After cancelling off the overall factor of $(2\pi)^4 \delta^{(4)}(p_1 + p_2)$ we find that the first diagram (from left to right) is equal to

$$D_1 = \left[\frac{1}{2} \right] \left(\frac{i}{p^2 - M^2} \right)^2 \int \frac{d^4 q}{(2\pi)^4} \frac{i}{q^2 - M^2} \frac{i}{(p - q)^2 - M^2} , \quad (4.34)$$

the second is

$$D_2 = \left[\frac{1}{2} \right] \left(\frac{i}{p^2 - M^2} \right)^2 \frac{i}{(0 - M^2)} \int \frac{d^4 q}{(2\pi)^4} \frac{i}{q^2 - M^2} , \quad (4.35)$$

and the third is given by

$$D_3 = \left[\frac{1}{4} \right] \left(\frac{i}{p^2 - M^2} \right)^2 \left[\int \frac{d^4 q}{(2\pi)^4} \frac{i}{q^2 - M^2} \right]^2 (2\pi)^4 \delta^{(4)}(p) . \quad (4.36)$$

Of the three, this last piece is the only disconnected one. In each case, the q integrations run over the momenta that are not completely fixed by momentum conservation. Applying our Feynman rules to the 3-point function at order- g^1 , it is straightforward to check that the results of Eqs. (4.23,4.24) are also reproduced.

4.2.2 Feynman Rules for Connected Amplitudes

With Feynman rules formulated in momentum space, it is now really easy to compute the connected amplitudes for scattering using the LSZ formula. The procedure to find the amplitude for $\langle \vec{p}_1 \vec{p}_2 \dots \vec{p}_n | \vec{k}_1 \dots \vec{k}_m \rangle_c$ at order g^M is:

1. Use the rules outlined above to find the perturbative value of $\tilde{G}^{(m+n)}(-p_1, -p_2, \dots, -p_n, k_1, k_2)$ keeping terms up to order g^M .
2. Remove all the diagrams that are not completely connected.
3. Cancel off the external propagator factors. Equivalently, multiply each diagram by $(-i)^{n+m} \prod_{i=1}^n (p_i^2 - M^2 + i\epsilon) \times \prod_{j=1}^m (k_j^2 - M^2 + i\epsilon)$.
4. Find \sqrt{Z} perturbatively by computing the 2-point function up to order g^M , multiply everything by $1/(\sqrt{Z})^{(n+m)}$, and keep terms only up to order g^M .
5. Take the limit $p_i^2 \rightarrow M^2$ for all external momenta.

The final result of all these steps is $\langle \vec{p}_1 \vec{p}_2 \dots \vec{p}_n | \vec{k}_1 \dots \vec{k}_m \rangle_c$

As a simple example, let us compute the connected amplitude for $\langle \vec{p}_2 \vec{p}_3 | \vec{p}_1 \rangle$ at order g^1 . The only connected part is the first term T_1 found in Eq. (4.23). Following our rules, the amplitude at this order is

$$\langle \vec{p}_2 \vec{p}_3 | \vec{p}_1 \rangle = (-ig) \times (2\pi)^4 \delta^{(4)}(p_2 + p_3 - p_1) . \quad (4.37)$$

This result explains why we assign a value of $-ig$ to each vertex.

The only slightly tricky part about these rules are the factors of \sqrt{Z} . Recall that this factor emerged in the Källén-Lehmann spectral representation in the form of Eqs. (4.17,4.19). They are called wavefunction renormalization factors, and they come from interactions changing the normalization of the field as defined by the residue of the one-particle pole of the 2-point function. In the present case, we have $Z = 1 + Ag^2 + \dots$, so we can take $Z = 1$ if we are working to leading non-trivial order in the coupling. This will be the case for most of what we will do, but we will come back later on in the course to discuss how to deal with the wavefunction factors properly.

4.3 Scattering and Decays

We now have all the tools we need to begin computing physical observables in the form of scattering cross sections and particle decay rates. These are some of the main observables that perturbative quantum field theory is used to calculate.

4.3.1 Scattering Cross Sections

A typical scattering experiment consists of an initial state of two well-separated particles that collide with each other to create a final state with n independent particles that propagate off to infinity (or thereabouts). Using perturbation theory, we can compute the vacuum matrix element $\langle \Omega | T \{ \phi(x_3) \dots \phi(x_{n+2}) \phi(z_1) \phi(z_2) \} | \Omega \rangle$. The LSZ formula then allows us to relate this

vacuum matrix element to the connected matrix element $\langle \vec{p}_2 \dots \vec{p}_{n+2} | \vec{k}_1 \vec{k}_2 \rangle_c$. This quantity is related to the scattering amplitude \mathcal{M} by

$$-i\mathcal{M} = \langle \vec{p}_3 \dots \vec{p}_{n+2} | \vec{k}_1 \vec{k}_2 \rangle_c \bigg/ (2\pi)^4 \delta^{(4)}(k_1 + k_2 - p_3 - \dots - p_{n+2}). \quad (4.38)$$

Since the connected matrix element is always proportional to the overall 4-momentum delta function, this equation just says that we should cancel off the delta function to get the amplitude.

The $2 \rightarrow n$ scattering cross section, corresponding to the total probability of non-trivial scattering per unit initial flux, is related to the scattering amplitude by

$$\sigma = \frac{S}{|\vec{v}_{rel}|} \frac{1}{2E_1 2E_2} \int \frac{d^3 p_3}{(2\pi)^3 2E_3} \dots \int \frac{d^3 p_{n+2}}{(2\pi)^3 2E_{n+2}} (2\pi)^4 \delta^{(4)}(k_1 + k_2 - \sum_{i=3}^{n+2} p_i) |\mathcal{M}|^2, \quad (4.39)$$

where $|\vec{v}_{rel}| = \sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2} / E_1 E_2$ is the magnitude of the initial relative velocity, and S is a combinatoric factor equal to one times $1/N!$ for every set of N identical particles in the final state. Derivations of this result can be found in Peskin & Schroeder [1], Srednicki [2], and Griffiths [10].

The result of Eq. (4.39) has a lot going on within it, but its physical content is very simple. First, $|\mathcal{M}|^2$ is the probability density for a single initial state $\vec{k}_1 + \vec{k}_2$ to scatter into the specific final state $\vec{p}_3 + \dots + \vec{p}_{n+2}$. The delta function enforces overall 4-momentum conservation. The scattering probability density is then summed over all distinct final states with a relativistic normalization. Collectively, this set of final states is often called the *phase space*. The prefactor before the integrations is a normalization to convert the result for a single initial state to the scattering probability rate per unit incident flux (= number of incident particles per unit area per unit time). At the end of the day, the cross section has units of area. The factor of S accounts for sets of indistinguishable particles.

e.g. 1. $2 \rightarrow 2$ Scattering in the $g\phi^3/3!$ Theory

To compute the scattering amplitude, we need to draw all the Feynman diagrams for this theory that can contribute. For now, we will only compute the leading non-trivial contribution in g . This comes from the diagrams in Fig. 4.2, which I have drawn with time going from left to right, and p_1 and p_2 as the initial momenta and p_3 and p_4 as the final momenta. The amplitude at this order is

$$-i\mathcal{M} = (-ig)^2 \left[\frac{i}{(p_1 + p_2)^2 - M^2} + \frac{i}{(p_1 - p_3)^2 - M^2} + \frac{i}{(p_1 - p_4)^2 - M^2} \right]. \quad (4.40)$$

These three diagrams are called *s*-, *t*-, and *u*-channel respectively. This designation corresponds to the three Lorentz invariant combinations of momenta relevant for $2 \rightarrow 2$ scattering:

$$s = (p_1 + p_2)^2 = (p_3 + p_4)^2 \quad (4.41)$$

$$t = (p_1 - p_3)^2 = (p_2 - p_4)^2 \quad (4.42)$$

$$u = (p_1 - p_4)^2 = (p_2 - p_3)^2. \quad (4.43)$$

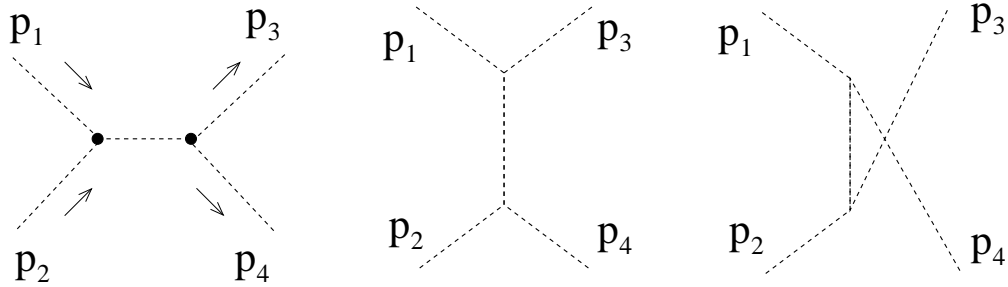


Figure 4.2: Feynman diagrams for *e.g.* 1.

Note that $s + t + u = 4M^2$. In terms of these combinations, we have

$$\mathcal{M} = g^2 \left[\frac{1}{s - M^2} + \frac{1}{t - M^2} + \frac{1}{u - M^2} \right]. \quad (4.44)$$

e.g. 2. Elastic Scattering in the CM Frame

Consider the $2 \rightarrow 2$ elastic scattering of a pair of particles of mass M . We will work in the center-of-mass (CM) frame where the sum of the initial 3-momenta vanishes. It is conventional to choose the z -axis in the direction of the incoming particles. Specifically,

$$p_1 = (E, 0, 0, p), \quad p_2 = (E, 0, 0, -p) . \quad (4.45)$$

with $E = \sqrt{p^2 + M^2}$. Applying the overall conservation of energy and momentum, the outgoing momenta must take the form

$$p_3 = (E, p \sin \theta, 0, p \cos \theta), \quad p_4 = (E, -p \sin \theta, 0, -p \cos \theta) , \quad (4.46)$$

where we have chosen to align the \vec{p}_3 axis to simplify the form of p_3 . In terms of these values, we can express the momentum invariants s , t , and u as

$$s = 4(M^2 + p^2), \quad t = -p^2(1 - \cos \theta)^2, \quad u = -p^2(1 + \cos \theta)^2 . \quad (4.47)$$

The scattering cross section in this frame is

$$\sigma = \frac{(1/2!)}{|\vec{v}_{rel}|} \frac{1}{4E^2} \int \frac{d^3 p_3}{2E_3(2\pi)^3} \int \frac{d^3 p_4}{2E_4(2\pi)^3} (2\pi)^4 \delta^{(4)}(p_1 + p_2 - p_3 - p_4) |\mathcal{M}|^2 \quad (4.48)$$

$$= \frac{1}{|\vec{v}_{rel}|} \frac{1}{32(2\pi)^2 E^2} \int_0^\infty dp' p'^2 \int d\Omega \frac{1}{E'^2} |\mathcal{M}|^2 \delta(2E - 2\sqrt{p'^2 + M^2}) \quad (4.49)$$

$$= \frac{1}{|\vec{v}_{rel}|} \frac{1}{64\pi E^2} \frac{p}{E} \left(\frac{1}{4\pi} \int d\Omega |\mathcal{M}|^2 \right) , \quad (4.50)$$

where $E' = \sqrt{p'^2 + m^2}$ and $|\vec{v}_{rel}| = 2\sqrt{1 - m^2/E^2}$ in this frame.

In many cases, the quantity of interest isn't the total cross section but rather the *differential cross section* per unit solid angle. To get this, just take the expression for the total cross section, but don't integrate over the outgoing particle direction you're interested in. For the $2 \rightarrow 2$ example here, we have

$$\frac{d\sigma}{d\Omega}(\theta, \phi) = \frac{1}{|\vec{v}_{rel}|} \frac{1}{256\pi^2 E^2} \frac{p}{E} |\mathcal{M}|^2 , \quad (4.51)$$

where $d\Omega = d(\cos \theta) d\phi$ is the solid angle corresponding to the direction of one of the outgoing particles.

4.3.2 Decays

The decays of an unstable particle are probabilistic, but are characterized by an average decay rate Γ . Specifically, given an initial sample of N_0 particles at time $t = 0$, the number of particles after time t is

$$N(t) = N_0 e^{-\Gamma t} . \quad (4.52)$$

The lifetime τ of a particle species is defined to be

$$\tau = 1/\Gamma . \quad (4.53)$$

Sometimes you will also hear of half-lives, given by $\tau_{1/2} = \tau \ln 2$. In natural units, the decay rate has units of mass.

The decay rate can be computed using the LSZ formula, even though this formula only really applies to stable particles that are able to propagate off to infinity. However, it turns out that the LSZ formula is also a good approximation for particles that are unstable but whose decay rates are very slow relative to their mass. In this case, the partial rate for an unstable particle of mass M at rest to decay to a final state containing n particles ($1 \rightarrow 2 + 3 + \dots + n + 1$) is

$$\Gamma(1 \rightarrow n) = \frac{S}{2M} \int \frac{d^3 p_2}{2E_2(2\pi)^3} \cdots \int \frac{d^3 p_{n+1}}{2E_{n+1}(2\pi)^3} (2\pi)^4 \delta^{(4)}(p_1 - \sum_{i=2}^{n+1} p_i) |\mathcal{M}|^2 , \quad (4.54)$$

where $|\mathcal{M}|^2$ is the corresponding $1 \rightarrow n$ amplitude defined in the same way as for scattering, and S is the symmetry factor. The total decay rate is the sum of the partial rates Γ_f of all the individual decay channels,

$$\Gamma = \sum_f \Gamma_f = \Gamma \sum_f BR_f , \quad (4.55)$$

where $BR_f = \Gamma_f/\Gamma$ is the *branching ratio* to the final state f .

Chapter 5

Poincaré and Particles

So far in the course we have glossed over Lorentz invariance and we have only studied theories that describe particles with spin equal to zero. In these notes we will tackle Lorentz and translation invariance head on, and this will lead us to theories with particles of non-trivial spin.

Before addressing Lorentz invariance, we will first discuss the general features of symmetries and how to implement them in quantum mechanics. With this in hand, we will focus on the Poincaré group consisting of Lorentz transformations plus translations and which is the symmetry group of flat spacetime. In particular, we will study the implications of Poincaré invariance on the structure of quantum fields and particle states in the Hilbert space.

5.1 Symmetries, Groups, and Representations

Symmetry transformations obey the mathematical properties of a *group*, and it is worth spending a bit of time discussing what they are.¹ A group G is a set of objects together with a multiplication rule such that:

1. if $f, g \in G$ then $h = f \cdot g \in G$ (closure)
2. $f \cdot (g \cdot h) = (f \cdot g) \cdot h$ (associativity)
3. there exists an identity element $1 \in G$ such that $1 \cdot f = f \cdot 1 = f$ for any $f \in G$ (identity)
4. for every $f \in G$ there exists an inverse element f^{-1} such that $f \cdot f^{-1} = f^{-1} \cdot f = 1$ (invertability)

Each group element corresponds to a different transformation of the same class.

¹ Much of this discussion is based on Refs. [11, 1], both of which provide a much more detailed account of the topics covered here.

A group can be defined via a multiplication table which specifies the value of $f \cdot g$ for every pair of elements $f, g \in G$. An *Abelian* group is one for which $f \cdot g = g \cdot f$ for every pair of $f, g \in G$. A familiar example of an Abelian group is the set of rotations in two dimensions. In contrast, the set of rotations in three dimensions is non-Abelian.

5.1.1 Representations of Groups

For the most part, we will be interested in symmetry transformations that act linearly on states in a Hilbert space,

$$|v_i\rangle \rightarrow |v'_i\rangle = U_{ij}|v_j\rangle. \quad (5.1)$$

As a result, we will usually work with matrix *representations* of groups. Groups themselves are abstract mathematical objects. A representation of a group is a set of $n \times n$ matrices $U(g)$, one for each group element, such that:

1. $U(f)U(g) = U(f \cdot g)$
2. $U(1) = \mathbb{I}$, the identity matrix.

Note that these conditions imply $U(f^{-1}) = U^{-1}(f)$. The value of n is called the dimension of the representation. For any group, there is always the *trivial* representation where $U(g) = \mathbb{I}$ for every $f \in G$. Note that a representation does not have to faithfully reproduce the full multiplication table. A representation is said to be *unitary* if all the representation matrices can be taken to be unitary ($U^\dagger = U^{-1}$).

e.g. 1. Rotations in two dimensions

This group is formally called $SO(2)$ and can be defined as an abstract mathematical object. Any group element can be associated with a rotation angle θ . The most familiar representation is in terms of 2×2 matrices,

$$D(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}. \quad (5.2)$$

Of course, there is also the trivial representation.

Our focus will be primarily on continuous transformations. These correspond to what are called *Lie groups*, which are simply groups whose elements can be parametrized in terms of a set of continuous variables $\{\alpha^a\}$. We can (and will) always choose these coordinates (near the identity) such that the point $\alpha^a = 0$ corresponds to the identity element of the group. Thus, for any representation of the group, we have for infinitesimal transformations near the identity

$$U(\alpha^a) = \mathbb{I} - i\alpha^a t^a + \mathcal{O}(\alpha^2). \quad (5.3)$$

The matrices t^a are called *generators* of the representation. Finite transformations can be built up from infinitesimal ones according to

$$U(\alpha^a) = \lim_{p \rightarrow \infty} (1 - i\alpha^a t^a / p)^p = e^{-i\alpha^a t^a}. \quad (5.4)$$

This is nice because it implies that we only need to sort out a finite set of generators when discussing the representation of a Lie group rather than the infinite number of group elements. Let us also mention that a Lie group is said to be *compact* iff all the parameters α^a run over finite intervals.

A set of generator matrices $\{t^a\}$ can represent a Lie group provided they satisfy a *Lie algebra*. Besides being able to add and multiply them, they must also satisfy the following conditions:

1. $[t^a, t^b] = if^{abc}t^c$ for some constants f^{abc}
2. $[t^a, [t^b, t^c]] + [t^b, [t^c, t^a]] + [t^c, [t^a, t^b]] = 0$ (Jacobi Identity)

The first condition is needed for the closure of the group (*i.e.* $\exp(-i\alpha^a t^a) \exp(-i\beta^a t^a) = \exp(-i\lambda^a t^a)$ for some λ^a) while the second is required for associativity. In fact, we can define a Lie group abstractly by specifying the *structure constants* f^{abc} . Most of the representations we'll work with are unitary, in which case the structure constants are all real and the generators t^a are Hermitian.

***e.g.* 2 $SU(2)$**

This is the prototypical Lie group, and it should already be familiar from what you know about spin in quantum mechanics. As a group, it is defined to be the set of 2×2 unitary matrices with determinant equal to one. The corresponding Lie algebra has three basis elements which satisfy

$$[t^a, t^b] = i\epsilon^{abc}t^c \quad (5.5)$$

The basic *fundamental* representation of $SU(2)$ is in terms of Pauli matrices: $t^a = \sigma^a/2$. Since $[\sigma^a, \sigma^b] = 2i\epsilon^{abc}\sigma^c$, it's clear that this is a valid representation of the algebra. You might also recall that any $SU(2)$ matrix can be written in the form $U = \exp(-i\alpha^a \sigma^a/2)$.

Even though $SU(2)$ is a group defined in terms of 2×2 matrices, it has other representations. You may recall from basic quantum mechanics that spin corresponds to a symmetry under $SU(2)$. The fundamental rep corresponds to $s = 1/2$, and $s = 0$ is just the trivial representation ($U(g) = \mathbb{I}$). On the other hand, we also know that there are spins with $s = 0, 1/2, 1, 3/2, \dots$, and these correspond to representations of dimension $(2s + 1)$.

5.1.2 More on Representations (optional)

The nice thing about working with linear generators t^a is that we can choose a nice basis for them. This is equivalent to choosing a nice set of coordinates for the Lie group. In

particular, it is always possible to choose the generators t_r^a of any representation r such that

$$\text{tr}(t_r^a t_r^b) = T_2(r) \delta^{ab}. \quad (5.6)$$

The constant $T_2(r)$ is called the Dynkin index of the representation. With the exception of the Poincaré group (Lorentz and translations), we will always implicitly work in bases satisfying Eq. (5.6), and we will concentrate on the case where the index is strictly positive. If so, the corresponding Lie group is said to be *compact* and is guaranteed to have finite-dimensional unitary representations. **This is not true of the Poincaré group, which is not compact.**

It turns out that there are only a finite set of classes of compact Lie groups. The *classical* groups are:

- $U(1)$ = phase transformations, $U = e^{i\alpha}$
- $SU(N)$ = set of $N \times N$ unitary matrices with $\det(U) = 1$
- $SO(N)$ = set of orthogonal $N \times N$ matrices with $\det(U) = 1$
- $Sp(2N)$ = set of $2N \times 2N$ matrices that preserve a slightly funny inner product.

In addition to these, there are the *exceptional* Lie groups: E_6, E_7, E_8, F_4, G_2 . In studying the Standard Model, we will focus primarily on $U(1)$ and $SU(N)$ groups.

Some useful and fun facts about compact Lie algebras:

- Except for $U(1)$, we have $\text{tr}(t^a) = 0$ for all the classical and exceptional Lie groups.
- Number of generators = $d(G)$

$$d(G) = \begin{cases} N^2 - 1; & SU(N) \\ N(N - 1)/2; & SO(N) \\ 2N(2N + 1)/2; & Sp(2N) \end{cases} \quad (5.7)$$

- A representation (= rep) is *irreducible* if it cannot be decomposed into a set of smaller reps. This is true if and only if it is impossible to simultaneously block-diagonalize all the generators of the rep. Irreducible representation = irrep.
- If one of the generators commutes with all the others, it generates a $U(1)$ subgroup called an Abelian factor: $G = G' \times U(1)$.
- If the algebra cannot be split into sets of mutually commuting generators it is said to be *simple*. For example, $SU(5)$ is simple (as are all the classical and exceptional Lie groups given above) while $SU(3) \times SU(2) \times U(1)$ is not simple. In the latter case, all the $SU(3)$ generators commute with all the $SU(2)$ generators and so on.

- A group is *semi-simple* if it does not have any Abelian factors.
- With the basis choice yielding Eq. (5.6), one can show that the structure constants are completely anti-symmetric.
- The *fundamental* representation of $SU(N)$ is the set of $N \times N$ special unitary matrices acting on a complex vector space. This is often called the \mathbf{N} representation. Similarly, the fundamental representation of $SO(N)$ is the set of $N \times N$ special orthogonal matrices acting on a real vector space.
- The *adjoint* ($= A$) representation can be defined in terms of the structure constants according to

$$(t_A^a)_{bc} = -if^{abc} \quad (5.8)$$

Note that on the left side, a labels the adjoint generator while b and c label its matrix indices.

- Given any rep t_r^a , the conjugate matrices $-(t_r^a)^*$ give another representation, unsurprisingly called the conjugate representation. A rep is said to be real if it unitarily equivalent to its conjugate. The adjoint rep is always real.
- The Casimir operator of a rep is defined by $T_r^2 = t_r^a t_r^a$ (with an implicit sum on a). One can show that T_r^2 commutes with all the t_r^a . For an irrep (=irreducible representation) of a simple group, this implies that

$$T_r^2 = C_2(r)\mathbb{I}, \quad (5.9)$$

for some positive constant $C_2(r)$.

- It is conventional to fix the normalization of the fundamental of $SU(N)$ such that $T_2(\mathbf{N}) = 1/2$. Once this is done, it fixes the normalization of all the other irreps. In particular, it implies that for $SU(N)$, $C_2(\mathbf{N}) = (N^2 - 1)/2N$, $T_2(A) = N = C_2(A)$.

5.2 Symmetries in Physical Systems

We have already discussed what it means for a transformation to be a symmetry of a classical system. Having discussed group representations, we can now apply this knowledge to choose convenient sets of variables for both classical and quantum systems.

5.2.1 Symmetries in Classical Mechanics

Recall that for a continuous classical system, the condition for the transformation

$$\phi_i \rightarrow \phi'_i = f_i(\phi) \quad (5.10)$$

to be a symmetry was

$$S[\phi] \rightarrow S[\phi'] := S'[\phi] = S[\phi] . \quad (5.11)$$

This usually implies that the Lagrangian should be unchanged by the transformation as well. Note as well that this is the *active picture*, and there is a similar relation for the passive picture.

In most cases of interest, we are interested in symmetry transformations that act linearly on the field variables. In this case, given a system described by n fields $\{\phi_1, \phi_2, \dots, \phi_n\}$, we have for each element g of the symmetry group

$$\phi_A(x) \rightarrow \phi'_A(x) = M_A^B(g) \phi_B(x) . \quad (5.12)$$

The matrices $M_A^B(g)$ form a representation of the symmetry group, in the sense described above, and there is a matrix for each element of the symmetry group. Note that these matrices $M(g)$ must be invertible, but they do not have to be unitary. For the Lagrangian to be invariant under the symmetry, it should be built up from terms that are also invariant. This strongly limits what can appear in the Lagrangian.

e.g. 3. A theory with $SU(2)$ symmetry.

Consider the theory of two of complex fields ϕ_1 and ϕ_2 combined to form a single two-component field Φ given by

$$\Phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} . \quad (5.13)$$

We will take the Lagrangian for the theory to be

$$\mathcal{L} = (\partial^\mu \Phi^\dagger)(\partial_\mu \Phi) - m^2 \Phi^\dagger \Phi - \frac{\lambda}{4} (\Phi^\dagger \Phi)^2 \quad (5.14)$$

This theory is invariant under $SU(2)$ transformations

$$\Phi(x) \rightarrow \Phi'(x) = e^{-i\alpha^a t^a} \Phi , \quad (5.15)$$

where $t^a = \sigma^a/2$ are the Pauli matrices. In this case, $M(\alpha^a) = e^{i\alpha^a \sigma^a/2}$. This is only one of many different theories with a symmetry under $SU(2)$ transformations. For instance, we could have also constructed a theory using the triplet representation of $SU(2)$ where the t^a matrices are proportional to the matrices that arise when you construct the spin operators $S_x, S_y,$ and S_z on states of spin $s = 1$. (Note that the $SU(2)$ in this example has nothing to do with spin!)

5.2.2 Symmetries in Quantum Mechanics

In quantum mechanics we have states in a Hilbert space and operators that act upon them. Observables correspond to the eigenvalues of Hermitian operators, and the squared norms

of states are interpreted as probabilities. In general, a transformation upon a quantum mechanical system can be implemented by an operator acting on all states:

$$|\psi\rangle \rightarrow |\psi'\rangle = U|\psi\rangle , \quad (5.16)$$

for some operator U . By construction, this operator is linear.

The transformations that tend to be the most interesting are symmetries. We already discussed what it means for a transformation to be a symmetry of a classical system. In quantum mechanics, we demand that a symmetry (not) do two things. The first is that it should not alter the inner products of states (which are interpreted as probabilities),

$$\langle\psi_a|\psi_b\rangle = \langle\psi'_a|\psi'_b\rangle \quad (5.17)$$

$$= \langle\psi_a|U^\dagger U|\psi_b\rangle , \quad (5.18)$$

which implies that we need $U^\dagger U = \mathbb{I}$. Thus, the transformation should be implemented by a unitary operator.² We would also like the transformation to be consistent with time evolution. In the Schrödinger picture, this means specifically that

$$|\psi'(t)\rangle = e^{-iH(t-t_0)}|\psi'(t_0)\rangle \quad (5.19)$$

$$Ue^{-iH(t-t_0)}|\psi(t_0)\rangle = e^{-iH(t-t_0)}U|\psi(t_0)\rangle , \quad (5.20)$$

which implies that³

$$[H, U] = 0 , \quad (5.21)$$

Together, Eqs (5.18,5.21) are the conditions that a set of transformations must satisfy to be a symmetry in quantum mechanics.

In general, the set of all symmetry transformations on the system forms a group. For each element g in the group, there is an operator $U(g)$ that acts linearly on the vector space of states. Therefore they must be a linear representation of the group. It is often useful to choose a basis of states that transform as independent blocks under the symmetry. For example, with spin we decompose the Hilbert space into states with different spins.

If we apply this definition to a continuous symmetry, we immediately get a quantum version of Noether's theorem. A continuous symmetry corresponds to a Lie group, and thus we can write any group element in terms of the generators of the Lie algebra as

$$U(\alpha^a) = e^{-i\alpha^a t^a} , \quad (5.22)$$

where α^a are the continuous parameters that label the different elements of the group and now t^a is a quantum operator that represents the Lie algebra of the group ($[t^a, t^b] = if^{abc}t^c$). By assumption, $U(\alpha^a)$ commutes with the Hamiltonian for any value of α^a . This implies that

$$[H, t^a] = 0 . \quad (5.23)$$

² Actually, this is a bit too strict since equality of the inner products up to a phase would have been fine as well. However, we will concentrate on unitary transformations in this course.

³This assumes that U has no explicit time dependence. If it does, the condition is $-i\partial_t U + [H, U] = 0$.

In the Heisenberg picture, this implies that the generators are conserved in that they do not evolve in time.

Just like we had the Schrödinger and Heisenberg pictures of time evolution in quantum mechanics, or the active and passive pictures of symmetries in classical mechanics, we can also think of symmetries as acting on operators rather than on states. It turns out that this is usually the more convenient thing to do in quantum field theories. In this picture, a symmetry transformation corresponds to keeping all the states the same but transforming the operators according to

$$\mathcal{O} \rightarrow \mathcal{O}' = U^\dagger(g)\mathcal{O}U(g) , \quad (5.24)$$

where $U(g)$ is unitary with $[H, U(g)] = 0$ as before. For continuous symmetries parametrized by α^a , we can consider an infinitesimal transformation and define

$$\mathcal{O}' - \mathcal{O} = \alpha^a (\Delta\mathcal{O})^a \quad (5.25)$$

Expanding $U(\alpha^a)$ and matching up α^a factors, we find

$$(\Delta\mathcal{O})^a = i[t^a, \mathcal{O}] . \quad (5.26)$$

This implies that the action of the transformation on the operator is encoded in its commutator with the generators (analogous to the commutator with H giving the time evolution).

In quantum field theories, we will usually be interested in symmetries that act linearly on the fields. By this, we mean that for a system of n fields $\{\phi_1, \dots, \phi_n\}$ we have

$$\phi_A(x) \rightarrow \phi'_A(x) = U^\dagger(g)\phi_A(x)U(g) = M_A^B(g)\phi_B(x) , \quad (5.27)$$

where the matrices M form a representation of the symmetry group.⁴ Note that we have three levels of mathematical structure here. The first is the abstract structure of the symmetry group itself, which need not refer to any matrices at all. The second level is the representation of the group on the states of the Hilbert space by the unitary operators U . And the third level is the representation of the group by the M matrices acting on the space fields. These matrices do not have to be unitary.

It is worth checking that this picture is consistent. We have trivially that $U(1) = \mathbb{I}$ and $M_A^B(1) = \delta_A^B$, and associativity and invertability follow by assumption. The last thing to check is closure. Suppose we transform by the group element $f \cdot g$. This gives on the left-hand side of Eq. (5.27)

$$U^\dagger(f \cdot g)\phi_A(x)U(f \cdot g) = U^\dagger(g)U^\dagger(f)\phi_A(x)U(f)U(g) \quad (5.28)$$

$$= U^\dagger(g) [M_A^B(f)\phi_B(x)] U(g) \quad (5.29)$$

$$= M_A^B(f)M_B^C(g)\phi_C(x) . \quad (5.30)$$

On the right-hand side, we find

$$M_A^C(f \cdot g)\phi_C(x) = M_A^B(f)M_B^C(g)\phi_C(x) . \quad (5.31)$$

Thus, everything works out as it should.

⁴ If the transformation shifts $x \rightarrow x'$ as well, this result is modified to $\phi'(x') = U^\dagger\phi(x')U = M\phi(x)$.

5.3 The Poincaré Group: Lorentz plus Translations

The symmetries we are most interested in for relativistic theories are those of spacetime, namely translations and Lorentz transformations (boosts and rotations). Together, this group of symmetries is called the Poincaré group. When we get to finding representations of the Poincaré group on fields, we will see that they correspond to particles of different spins.

5.3.1 The Translation Group

The Poincaré group consists of translations plus Lorentz transformations. We have already discussed translations,

$$x^\mu \rightarrow x^\mu + a^\mu . \quad (5.32)$$

Translations form a group that can be parametrized by the four components of the translation vector a^μ . We can represent the group on the space of functions of x by the linear operator

$$U(a) = e^{ia^\nu(i\partial_\nu)} , \quad (5.33)$$

that produces

$$U(a)f(x) = \left[1 - a^\nu \partial_\nu + \frac{1}{2!}(-a^\nu \partial_\nu)^2 + \dots \right] f(x) = f(x - a) . \quad (5.34)$$

Specializing to infinitesimal translations, we find the generators of this representation to be

$$P_\mu = i\partial_\mu . \quad (5.35)$$

These satisfy the Lie algebra

$$[P^\mu, P^\nu] = 0 . \quad (5.36)$$

The group of translations is formally defined as the abstract Lie group spanned by the parameters $a^\mu \in (-\infty, \infty)$ with Eq. (5.36) as the Lie algebra. Our labelling of the generators by P^μ reflects our result that the conserved currents corresponding to invariance under translations are the energy and momentum operators.

5.3.2 The Lorentz Group

The second component of the Poincaré group are Lorentz transformations. These are defined as the set of linear transformations of spacetime that leave the Lorentz interval unchanged. Specifically, a Lorentz transformation is any real linear transformation Λ such that

$$x^\mu \rightarrow x'^\mu = \Lambda^\mu_\nu x^\nu , \quad (5.37)$$

with

$$\eta_{\mu\nu} dx'^{\mu} dx'^{\nu} = \eta_{\mu\nu} dx^{\mu} dx^{\nu} . \quad (5.38)$$

This implies that the transformation matrices must satisfy

$$\eta_{\mu\nu} = \eta_{\alpha\beta} \Lambda^{\alpha}_{\mu} \Lambda^{\beta}_{\nu} , \quad (5.39)$$

which is equivalent to the condition

$$(\Lambda^{-1})^{\mu}_{\nu} = \Lambda^{\mu}_{\nu} := \eta_{\nu\beta} \eta^{\mu\alpha} \Lambda^{\beta}_{\alpha} . \quad (5.40)$$

Besides just x^{μ} , we call any four-component object v^{μ} transforming this way under Lorentz a *four vector*:

$$v^{\mu} \rightarrow v'^{\mu} = \Lambda^{\mu}_{\nu} v^{\nu} . \quad (5.41)$$

The result of Eq. (5.40) implies that the dot product of any pair of four vectors v^{μ} and w^{μ} is unchanged by Lorentz transformations,

$$v \cdot w := \eta_{\mu\nu} v^{\mu} w^{\nu} = v' \cdot w' . \quad (5.42)$$

More generally, we define an $(n, 0)$ -index Lorentz tensor to be an object $T^{\mu_1 \mu_2 \dots \mu_n}$ that transforms as

$$T^{\mu_1 \dots \mu_n} \rightarrow T'^{\mu_1 \dots \mu_n} = \Lambda^{\mu_1}_{\nu_1} \dots \Lambda^{\mu_n}_{\nu_n} T^{\nu_1 \dots \nu_n} . \quad (5.43)$$

From this point of view, a vector is just a $(1, 0)$ tensor. We also define tensors with lowered indices by starting with a $(n, 0)$ tensor and lowering some of the indices with $\eta_{\mu\nu}$. We call a tensor with n upper indices and m lower indices a (n, m) tensor. It is straightforward to show that any product of tensors with all the indices contracted is Lorentz-invariant. More generally, any quantity defined on spacetime can be decomposed into Lorentz tensors as far as its Lorentz transformation properties are concerned.

Let us turn next to the Lie group structure of Lorentz transformations. Since this group is defined in terms of a set of linear transformations on 4-vectors, we already know one representation of the group. By matching these transformations to the general structure of Lie groups, we can figure out how to build other representations. The identity element of the group is clearly $\Lambda^{\mu}_{\nu} = \delta^{\mu}_{\nu}$. Expanding around the identity, we have

$$\Lambda^{\mu}_{\nu} = \delta^{\mu}_{\nu} + \omega^{\mu}_{\nu} + \dots \quad (5.44)$$

If we plug this into the requirement of Eq. (5.38), we find that the the only condition on ω^{μ}_{ν} is that it be antisymmetric:

$$\omega_{\mu\nu} = -\omega_{\nu\mu} . \quad (5.45)$$

With this constraint, ω^{μ}_{ν} has six independent elements that we can identify with the three generators of spatial rotations and the three generators of Lorentz boosts.

To put this another way, we have just seen that the Lorentz group can be parametrized in terms of six real numbers, which we can write as a antisymmetric (0,2) tensor $\omega_{\mu\nu}$. A general unitary representation of the Lorentz group must therefore take the form

$$U(\omega) = e^{-i\omega_{\mu\nu}J^{\mu\nu}/2} = \mathbb{I} - \frac{i}{2}\omega_{\mu\nu}J^{\mu\nu} + \dots, \quad (5.46)$$

where the $J^{\mu\nu}$ are a set of six antisymmetric Hermitian operators that generate the group representation.⁵ As an operator, consistency of the representation requires that $J^{\mu\nu}$ transforms as a (2,0) tensor,

$$U^\dagger(\Lambda)J^{\mu\nu}U(\Lambda) = \Lambda^\mu_\rho\Lambda^\nu_\sigma J^{\rho\sigma}. \quad (5.47)$$

Expanding this result out when $\Lambda^\mu_\nu = \delta^\mu_\nu + \omega^\mu_\nu$ is also infinitesimal gives the commutation relations of the $J^{\mu\nu}$ generators:

$$[J^{\mu\nu}, J^{\rho\sigma}] = -\frac{i}{2}(\eta^{\mu\rho}J^{\nu\sigma} - \eta^{\mu\sigma}J^{\nu\rho}) \quad (5.48)$$

$$= -i(\eta^{\mu\rho}J^{\nu\sigma} - \eta^{\nu\rho}J^{\mu\sigma} - \eta^{\mu\sigma}J^{\nu\rho} + \eta^{\nu\sigma}J^{\mu\rho}). \quad (5.49)$$

These relations define the abstract Lie algebra of the Lorentz group.

The six independent $J^{\mu\nu}$ generators can be rewritten in terms of the more familiar generators of rotations and boosts. Let us define

$$J^i = \frac{1}{2}\epsilon^{ijk}J^{jk}, \quad K^i = J^{0i}. \quad (5.50)$$

The commutation relations of Eq. (5.49) then imply

$$[J^i, J^j] = i\epsilon^{ijk}J^k, \quad (5.51)$$

$$[K^i, K^j] = -i\epsilon^{ijk}J^k, \quad (5.52)$$

$$[J^i, K^j] = i\epsilon^{ijk}K^k. \quad (5.53)$$

The first should be familiar, while the second and third are the generalizations to boosts.

With the Lie algebra of Lorentz in hand, we can try to find representations of it. In fact, we already have one, namely the Lorentz group as a set of linear operators acting on the space of four vectors. It is not hard to check that the generators of this representation are

$$(J_4^{\mu\nu})_{\alpha\beta} = i(\delta^\mu_\alpha\delta^\nu_\beta - \delta^\mu_\beta\delta^\nu_\alpha). \quad (5.54)$$

Here, μ and ν label which generator we want, while α and β label the elements of the representation matrix. Explicitly, we have

$$\Lambda^\alpha_\beta = \left(\exp \left[-\frac{i}{2}\omega_{\mu\nu}(J_4^{\mu\nu}) \right] \right)_\beta^\alpha = \delta^\alpha_\beta - \frac{i}{2}\omega_{\mu\nu}(J_4^{\mu\nu})^\alpha_\beta + \dots \quad (5.55)$$

$$= \delta^\alpha_\beta + \omega^\alpha_\beta + \dots, \quad (5.56)$$

⁵The factor of two here is just a matter of tradition.

just like Eq. (5.44). Note that for any Λ , there exists a unique ω . A second simple representation is the Lorentz group acting on the space of functions of x . The generators are

$$J_x^{\mu\nu} = i(x^\mu \partial^\nu - x^\nu \partial^\mu) . \quad (5.57)$$

These generalize the familiar angular momentum operators $\vec{J} = \vec{x} \times (-i\vec{\nabla})$ in position-space quantum mechanics.

There's a low-down dirty trick to find the general representations of the Lorentz group. Let us define new operators A^i and B^i by

$$A^i = \frac{1}{2}(J^i + iK^i) , \quad B^i = \frac{1}{2}(J^i - iK^i) . \quad (5.58)$$

These operators have the commutation relations

$$[A^i, A^j] = i\epsilon^{ijk} A^k \quad (5.59)$$

$$[B^i, B^j] = i\epsilon^{ijk} B^k \quad (5.60)$$

$$[A^i, B^j] = 0 . \quad (5.61)$$

Note as well that $(A^i)^\dagger = B^i$ and so on. These factors each have the commutation relations of two independent $SU(2)$ Lie algebras. Thus, to find representations of Lorentz, all we need to do is specify the representations of each of the $SU(2)$ factors. The irreducible representations are therefore (j_A, j_B) , with j_A and j_B half-integer, and have dimension $(2j_A + 1) \times (2j_B + 1)$.

The lowest representation is just the trivial $(0, 0)$. It corresponds to the scalars we have been studying so far. The next two up are $(1/2, 0)$ and $(0, 1/2)$, both of dimension two. We will see that these correspond to left- and right-handed *Weyl fermions*. Under the rotation subgroup of Lorentz, they both unsurprisingly have spin $j = 1/2$. The next representation is the four-dimensional $(1/2, 1/2)$. It corresponds to a Lorentz vector, and it decomposes into spins $j = 0, 1$ under the rotation subgroup. Going to higher values of j_A and j_B gives even higher spins.

In the discussion so far, we have implicitly assumed that the elements of the Lorentz group can be continuously deformed to the identity. This isn't necessarily true. Two matrices that satisfy the condition of Eq. (5.39) that cannot be deformed smoothly to unity are

$$\mathcal{P} = \text{diag}(+1, -1, -1, -1) , \quad (5.62)$$

and

$$\mathcal{T} = \text{diag}(-1, +1, +1, +1) . \quad (5.63)$$

The matrix \mathcal{P} is called a parity transformation and \mathcal{T} is called time reversal. It turns out that any Lorentz transformation can be written as the product of a transformation connected to the identity times either \mathcal{P} , \mathcal{T} , or \mathcal{PT} . In this sense, the Lorentz group has four independent sectors. We will be interested mainly in the sector connected to the identity, called the *proper orthochronous* subgroup of Lorentz, which is often what is meant by "Lorentz group", and

from here on we will follow this convention. In Nature, it turns out that “Lorentz” is a good symmetry but parity and time-reversal are not.

Applying \mathcal{P} to the generators, we find that $J^i \rightarrow J^i$ and $K^i \rightarrow -K^i$. Correspondingly, we find $A^i \leftrightarrow B^i$. Thus, the effect of parity is to map $(j_A, j_B) \leftrightarrow (j_B, j_A)$. The Weyl spinor irreps of Lorentz are therefore not representations of the Lorentz group extended by parity. On the other hand, the reducible representation $(1/2, 0) \oplus (0, 1/2)$ does work. We will use this rep soon to describe electrons in QED.

5.3.3 The Poincaré Group

The full Poincaré group consists of translations and Lorentz, and the general form of such a transformation is

$$x^\mu \rightarrow \Lambda^\mu_\nu x^\nu + a^\mu . \quad (5.64)$$

Thus, we will label general Poincaré group elements by

$$\{\Lambda, a\} . \quad (5.65)$$

We would like to find representations of this larger group. To do so, we need to figure out the Lie algebra.

Since we already have the Lie algebras for the translation and Lorentz subgroups, the only other things we need are the commutation relations between P^μ and $J^{\rho\sigma}$, provided they don't induce any new operators. These can be obtained by applying a Lorentz transformation to the P^μ operator, which must transform like a 4-vector,

$$U^\dagger(\Lambda)P^\mu U(\Lambda) = \Lambda^\mu_\nu P^\nu . \quad (5.66)$$

Specializing to $\Lambda^\mu_\nu = \delta^\mu_\nu + \omega^\mu_\nu + \dots$ and $U(\Lambda) = \mathbb{I} - i\omega_{\rho\sigma}J^{\rho\sigma}/2 + \dots$, we find

$$[P^\mu, J^{\rho\sigma}] = i(\eta^{\mu\rho}P^\sigma - \eta^{\mu\sigma}P^\rho) . \quad (5.67)$$

Collecting this and our previous results, the full set of commutation relations for the Poincaré group is therefore

$$[P^\mu, P^\nu] = 0 \quad (5.68)$$

$$[P^\mu, J^{\rho\sigma}] = i(\eta^{\mu\rho}P^\sigma - \eta^{\mu\sigma}P^\rho) \quad (5.69)$$

$$[J^{\mu\nu}, J^{\rho\sigma}] = -i(\eta^{\mu\rho}J^{\nu\sigma} - \eta^{\nu\rho}J^{\mu\sigma} - \eta^{\mu\sigma}J^{\nu\rho} + \eta^{\nu\sigma}J^{\mu\rho}) , \quad (5.70)$$

Together, these relations completely fix the Lie algebra structure of the Poincaré group.

Looking at Eq. (5.69), it would appear that some of the $J^{\mu\nu}$ currents are not conserved, in that $[P^0, J^{\mu\nu}] \neq 0$ (recall that $P^0 = H$). Working out the details, the non-zero components correspond to the boosts,

$$[H, J^{0i}] = iP^i . \quad (5.71)$$

This contradicts Eq. (5.21), one of our conditions for an operator to be the generator of a symmetry. It turns out that Eq. (5.21) does not quite apply to the case where the operator has an explicit dependence on t (*i.e.* t appears in the operator on its own, and not as the argument of a field). The correct generalization is

$$0 = i[H, U] + \frac{\partial U}{\partial t} = \frac{dU}{dt} . \quad (5.72)$$

Working out the Noether current for $J^{\mu\nu}$ in a general Lorentz-invariant field theory, one finds that the J^{0i} components depend on t explicitly and satisfy the general condition of Eq. (5.72) [12].

5.3.4 Representations of Poincaré on Fields

To construct a Poincaré-invariant quantum field theory, it is convenient to use field variables that have well-defined transformation properties, corresponding to fields that transform under definite representations of the Poincaré group. Note that unlike the representations of groups on quantum mechanical states, the representations of Poincaré on fields do not have to be unitary.

Given a Poincaré transformation

$$x \rightarrow x' = \Lambda x + a , \quad (5.73)$$

the corresponding transformation on a field operator will be

$$\phi_A(x) \rightarrow \phi'_A(x) = U^\dagger(\Lambda, a)\phi_A(x)U^\dagger(\Lambda, a) , \quad (5.74)$$

with the transformation operator

$$U(\Lambda, a) = \mathbb{I} - \frac{i}{2}\omega_{\mu\nu}J^{\mu\nu} + ia_\mu P^\mu + \dots \quad (5.75)$$

Our challenge is to figure out what $\phi'_A(x)$ can be.

Specializing to the translation subgroup of Poincaré, we already know the result. Recall that under $x \rightarrow x' = x + a$, we had

$$\phi_A(x) \rightarrow \phi'_A(x) = \phi_A(x - a) = e^{-iP \cdot a}\phi_A(x)e^{iP \cdot a} . \quad (5.76)$$

This has precisely the form of Eq. (5.74) with

$$U(1, a) = e^{ia_\mu P^\mu} , \quad (5.77)$$

where now the Hermitian P^μ operators act on the Hilbert space.

Let us turn next to the Lorentz subgroup. The most simple representation of Lorentz is the trivial one, with $\phi(x) \rightarrow \phi'(x') = \phi(x)$. This implies

$$\phi(x) \rightarrow \phi'(x) = U^\dagger(\Lambda, 0)\phi(x')U(\Lambda, 0) = \phi(\Lambda^{-1}x) . \quad (5.78)$$

Put another way, the shifted field is the same function of the rotated coordinates as the original field on the original coordinates. For example, if $\phi(x)$ is zero everywhere except at $x = x_0$, $\phi'(x)$ will be zero everywhere except at $x'_0 = \Lambda x_0$ with $\phi'(\Lambda x_0) = \phi(x_0)$. This has the form of Eq. (5.27), but with an additional change in the x coordinate. We found previously that this trivial representation corresponds to particles of spin $s = 0$.

A simple generalization of this result is the Lorentz vector field A^μ . Here, the four component fields rotate into each other under a Lorentz transformation. More precisely,

$$A^\mu(x) \rightarrow A'^\mu(x) = U^\dagger(\Lambda, 0)A^\mu(x)U(\Lambda, 0) = \Lambda^\mu_\nu A^\nu(\Lambda^{-1}x) \quad (5.79)$$

We see that the field components are rotated in addition to the shift in the coordinate dependence. They clearly transform in the vector representation of Lorentz. This also has the form of Eq. (5.27) with the representation matrix $M_A^B \sim \Lambda^\mu_\nu$.

It is instructive to look at the infinitesimal form of this transformation, with $\Lambda = 1 + \omega$. Expanding the left side of Eq. (5.79) we find

$$U^\dagger(1 + \omega, 0)A^\mu(x)U(1 + \omega, 0) = A^\mu(x) + \frac{i}{2}\omega_{\alpha\beta}[J^{\alpha\beta}, A^\mu(x)] . \quad (5.80)$$

On the right side of Eq. (5.79), we find

$$\Lambda^\mu_\nu A^\nu(\Lambda^{-1}x) = A^\mu(x) - \frac{i}{2}\omega_{\alpha\beta} [i(x^\alpha \partial^\beta - x^\beta \partial^\alpha)A^\mu(x) + i(\eta^{\mu\alpha} \delta^\beta_\nu - \eta^{\mu\beta} \delta^\alpha_\nu)A^\nu(x)] . \quad (5.81)$$

Comparing both sides, we see that

$$i[J^{\alpha\beta}, A^\mu(x)] = -iJ_x^{\alpha\beta} A^\mu(x) - i(J_4^{\alpha\beta})^\mu_\nu A^\nu(x) . \quad (5.82)$$

Comparing to the general result of Eq. (5.26), the left side just says that this is the variation of the $A^\mu(x)$ operator under infinitesimal Lorentz transformations. On the right side, there are two contributions. The first term is the variation due to the shift in the spacetime coordinate x . The second term is due to the vector transformation of the field components. Specializing to the ij components, this is just the expression for the infinitesimal change in the vector field under spatial rotations. In this case, the first term corresponds to the orbital angular momentum carried by the field while the second is due to the intrinsic ‘‘spin’’ of the field. Note that there is no spin contribution for $\mu = 0$.

Given this result for the vector field, the generalization to an arbitrary representation of the Lorentz subgroup acting on the fields $\{\phi_C\}$ should not be too surprising. It is

$$i[J^{\alpha\beta}, \phi_C(x)] = -iJ_x^{\alpha\beta} \phi_C(x) - i(J_{AB}^{\alpha\beta})_C^D \phi_D(x) , \quad (5.83)$$

where $J_{AB}^{\alpha\beta}$ is the generator of the (j_A, j_B) representation of the Lorentz Lie algebra discussed in Section 5.3.2. This relation gives the infinitesimal variation in the field $\phi_C(x)$ under Lorentz. Composing them, we can build up the general form of a finite Lorentz transformation:

$$U^\dagger(\Lambda, 0)\phi_C(x)U(\Lambda, 0) = \left[\exp\left(-\frac{i}{2}\omega_{\alpha\beta} J_{AB}^{\alpha\beta}\right) \right]_C^D \phi_D(\Lambda^{-1}x) . \quad (5.84)$$

We can now combine translations and Lorentz transformations and give the general form a finite Poincaré transformation on a field. It is just

$$U^\dagger(\Lambda, a)\phi_C(x)U(\Lambda, a) = \left[\exp\left(-\frac{i}{2}\omega_{\alpha\beta}J_{AB}^{\alpha\beta}\right) \right]_C^D \phi_D(\Lambda^{-1}x - a) . \quad (5.85)$$

The transformation matrix on the right-hand side describes the representation of the field under Lorentz, while the change in the coordinate of the field contains the rest of Lorentz together with all of Poincaré.

5.3.5 Representations of Poincaré on States

Finding the representations of the Poincaré group on one-particle quantum states is slightly more challenging. The difficulty lies in the fact that we are only allowed to use unitary representations. This has a very important consequence: physical particle states with mass $m > 0$ can be characterized completely by their mass and their representation under the $SU(2)$ spin group. In contrast, massless particles states can be characterized by their representation under a slightly different group, called helicity. For more details, see Refs. [13, 3, 14].

The translation part of the Poincaré group is trivial to include if we work with simultaneous eigenstates of P^μ :

$$P^\mu|p, \sigma\rangle = p^\mu|p, \sigma\rangle , \quad (5.86)$$

where the σ index labels any other properties the state could have under Lorentz. Applying a Lorentz transformation to such a state, we must have

$$U(\Lambda)|p, \sigma\rangle = [D_p(\Lambda)]_{\sigma\sigma'}|\Lambda p, \sigma'\rangle , \quad (5.87)$$

where $[D_p]_{\sigma\sigma'}$ is a matrix describing how states transform into each other under Lorentz. Our remaining task is to figure out what these matrices can be.

Since we are using momentum eigenstates, these states are also eigenstates of the $P^2 = P_\mu P^\mu$ operator. This operator is invariant under both translations and Lorentz, which can be verified by checking that it commutes with all the Poincaré generators P^μ and $J^{\mu\nu}$. Thus, for a set of states transforming under a given representation of Poincaré, they must all have the same eigenvalue of P^2 . For this reason, P^2 is called a *Casimir invariant* of the group and its eigenvalues are usually labelled by $P^2 = M^2$. If we also use the fact that (proper orthochronous) Poincaré transformations do not change the sign of p^0 , we can subdivide the possible representations into nine subclasses according to whether M^2 and p^0 are positive, negative, or zero. The cases of physical relevance are $M^2 = 0 = p^0$, $M^2 > 0$ and $p^0 > 0$, and $M^2 = 0$ and $p^0 > 0$. The first of these corresponds to the vacuum $|\Omega\rangle$ which transforms under the trivial representation. The second and third correspond to massive and massless particles.

For a massive particle with mass M , let us define a reference momentum

$$k^\mu = (M, \vec{0}) , \quad (5.88)$$

corresponding to the rest frame of the particle. We can get to any other momentum p (with $p^2 = M^2$) in a unique way by applying the Lorentz transformation L_p . On states, this implies that

$$|p, \sigma\rangle = U(L_p)|k, \sigma\rangle . \quad (5.89)$$

Note that the same index σ appears on both sides of the equation. This corresponds to a specific choice for how to relate the index of the reference state to the index of the more general state.

For arbitrary transformations, the transformed state will have a modified index structure as in Eq. (5.87). Performing another Lorentz transformation Λ , we get (leaving out the σ stuff for now)

$$U(\Lambda)|p\rangle = U(\Lambda)U(L_p)|k\rangle \quad (5.90)$$

$$= U(L_{\Lambda p})U^\dagger(L_{\Lambda p})U(\Lambda)U(L_p)|k\rangle \quad (5.91)$$

$$= U(L_{\Lambda p})U(L_{\Lambda p}^{-1}\Lambda L_p)|k\rangle . \quad (5.92)$$

This simple rearrangement has an important physical implication. The argument of the second U operator in the last line is a Lorentz transformation that takes $k \rightarrow p \rightarrow \Lambda p \rightarrow k$. It is therefore an element of the *little group* of Lorentz, the subgroup of Lorentz that maps k^μ to itself. To figure out the effect of the transformation on the σ indices, it is therefore sufficient to find the representations of the little group alone.

Given the form of the reference momentum in Eq. (5.88), the little group for the case of a massive particle is just the spatial rotation subgroup of Lorentz generated by the J^a , $a = 1, 2, 3$. This subgroup is just $SU(2)$, and we can therefore identify the σ label with a spin index! The unitary one-particle representations of a massive particle are therefore labelled by its 4-momentum and spin. This might not seem too surprising, but here we have found it to be a consequence of the underlying Poincaré symmetry.

We also know how to build representations of $SU(2)$, and this allows us to deduce the form of the $[D_p(\Lambda)]$ matrices. For any element W^μ_ν of the little group ($W^\mu_\nu k^\nu = k^\mu$), we have

$$U(W)|k, \sigma\rangle := [d(W)]_{\sigma\sigma'}|k, \sigma'\rangle , \quad (5.93)$$

where the little group matrix is

$$d(W) = e^{-i\theta^a J^a_j} , \quad (5.94)$$

for some spin- j representation J^a_j of $SU(2)$. The generators therefore have dimension $(2j + 1) \times (2j + 1)$ and the σ index can be identified with m_j values (for states $|j, m_j\rangle$).

Applying this to a general Lorentz transformation on a general state,

$$U(\Lambda)|p, \sigma\rangle = U(L_{\Lambda p})U(W(\Lambda, p))|k, \sigma\rangle \quad (5.95)$$

$$[D_p(\Lambda)]_{\sigma\sigma'}|\Lambda p, \sigma'\rangle = U(L_{\Lambda p})[d(W(\Lambda, p))_{\sigma\sigma'}|k, \sigma'\rangle \quad (5.96)$$

$$= [d(W(\Lambda, p))_{\sigma\sigma'}|\Lambda p, \sigma'\rangle . \quad (5.97)$$

Comparing both sides, we see that $D_p(\Lambda) = d(W(\Lambda, p))$, where $W(\Lambda, p)$ is the combined little group Lorentz transformation that sends $k \rightarrow p \rightarrow \Lambda p \rightarrow k$.

The total spin of a representation can be related to a second Casimir operator. Let us define the *Pauli-Lubanski pseudovector* by

$$W_\mu = -\frac{1}{2}\epsilon_{\mu\nu\rho\sigma}J^{\nu\rho}P^\sigma, \quad (5.98)$$

where $\epsilon^{\mu\nu\rho\sigma}$ is the totally antisymmetric tensor with $\epsilon^{0123} = +1$. Since $W_\mu P^\mu = 0$, in the particle rest frame we have $W^\mu = (0, \vec{W})$ with

$$W^i = M J^i. \quad (5.99)$$

The quantity $W^2 = W_\mu W^\mu$ is therefore Lorentz invariant and equal to

$$W^2 = -M^2 \vec{J} \cdot \vec{J} \rightarrow -M^2 s(s+1). \quad (5.100)$$

With a bit of work, you can show that W^2 also commutes with all the Poincaré generators. Thus, it is a Casimir invariant of the group and is equal to a fixed number in any representation. In all, we can completely characterize any massive representation of Poincaré by the mass M and total spin s .

These arguments go through in the same way in the massless case with one important difference; a massless particle has no rest frame. Instead, a reasonable choice for the reference momentum is

$$k^\mu = (k, 0, 0, k). \quad (5.101)$$

The little group is no longer the rotation group. Instead, it is generated by J^3 , $L_1 = K^1 - J^2$, and $L_2 = K^2 + J^1$, which have the Lie algebra of a different group called $ISO(2)$. It turns out that the only finite-dimensional representation of this group has $W^2 = 0$ along with $P^2 = 0$. This implies that

$$W^\mu |\vec{k}\rangle = h P^\mu |\vec{k}\rangle \quad (5.102)$$

on any such massless state. The proportionality constant is called the *helicity* of the particle, and is equal to

$$h = \frac{\vec{J} \cdot \vec{p}}{|\vec{p}|} = W^0 / |\vec{p}|. \quad (5.103)$$

This single number characterizes the finite representation of the little group, and must either be integer or half-integer. It turns out that the consistency of a theory requires that it contain states with both positive and negative helicity. These two helicity eigenstates correspond to “spin” parallel ($h > 0$) and antiparallel ($h < 0$) to the direction of motion. This result has important implications for the photon. First, it agrees with the fact that the photon only has two independent polarizations. Second, we will see that the one-particle state of the photon will be created by a vector field A^μ with four components, only two of which can correspond to physical excitations. Getting rid of the extra degrees of freedom will turn out to be highly non-trivial and will lead us to gauge invariance.

5.4 Executive Summary

Here's the short version of all this:

- Symmetry transformations have the mathematical structure of a group.
- A representation of a group is a set of linear operators $M(g)$ (which can be written as matrices once we specify a basis) that obey the group multiplication rules: $M(1) = \mathbb{I}$ and $M(f \cdot g) = M(f)M(g)$. Even though the representation is properly the set of linear operators, sometimes the vector space upon which they act is also called “the representation” of the group.
- A Lie group is one that can be parametrized by continuous coordinates $\{\alpha^a\}$. For group elements connected to the identity, we have $U(\alpha^a) = \exp(-i\alpha^a t^a)$. To represent such elements, we only need to find a representation of the Lie algebra for t^a .
- In quantum mechanics, symmetries are implemented by unitary operators $U(g)$ acting on states with $[H, U] = 0$. Equivalently, we can transform the operators instead of the states according to $\mathcal{O} = U^\dagger(g)\mathcal{O}U(g)$. For Lie groups, this implies that $\partial_t t^a = 0$ is conserved and $(\Delta\mathcal{O})^a = i[t^a, \mathcal{O}]$.
- Poincaré = translations plus Lorentz. The corresponding Lie algebra is

$$[P^\mu, P^\nu] = 0 \quad (5.104)$$

$$[P^\mu, J^{\rho\sigma}] = i(\eta^{\mu\rho}P^\sigma - \eta^{\mu\sigma}P^\rho) \quad (5.105)$$

$$[J^{\mu\nu}, J^{\rho\sigma}] = -i(\eta^{\mu\rho}J^{\nu\sigma} - \eta^{\nu\rho}J^{\mu\sigma} - \eta^{\mu\sigma}J^{\nu\rho} + \eta^{\nu\sigma}J^{\mu\rho}) . \quad (5.106)$$

Translations are easy to represent using functions of spacetime. For Lorentz, we can rewrite the generators as a pair of $SU(2)$ factors, and the representations are labelled by (j_A, j_B) with $j_{A,B} = 0, 1/2, 1, 3/2, \dots$

- We will use quantum fields that transform under definite representations of Poincaré. This means that as a quantum operator

$$U^\dagger(\Lambda, a)\phi_A(x)U(\Lambda, a) = M_A^B(\Lambda)\phi_B(\Lambda^{-1}x - a) , \quad (5.107)$$

where the matrices $M_A^B(\Lambda)$ form a (possibly non-unitary) representation of Lorentz.

- Representations of Poincaré in terms of the U operators acting on states is more complicated because they have to be unitary. Any representation can be characterized by the values of the Casimir operators P^2 and W^2 that commute with all the generators. It turns out that for momentum eigenstates, we only need to find representations of the much simpler little group. For massive states, $P^2 = M^2 > 0$, the little group is $SU(2)$ and they can be labelled by their momentum and their spin. For massless states, we have a slightly different little group $ISO(2)$ and the states are labelled by momentum and helicity.

Chapter 6

Fun with Fermions

Having established how fields transform under Poincaré, we turn next to the simplest non-trivial representations. These are the two-component representations $(j_A, j_B) = (1/2, 0)$, $(0, 1/2)$, that are usually referred to as left- and right-handed Weyl fermions. When we get around to quantizing the theory later, we will see that they do indeed correspond to fermionic particles. We will also study the larger reducible representation $(1/2, 0) \oplus (0, 1/2)$, which corresponds to a four-component object called a Dirac fermion. For more details, see Refs. [15, 16, 17, 1, 2, 3].

6.1 Weyl Fermions

A left-handed Weyl fermion is an object transforming under the $(1/2, 0)$ representation of the Lorentz group. Similarly, a right-handed Weyl fermion is an object transforming under the $(0, 1/2)$ rep. Both are two-component objects that we call *spinors*. In describing their transformation properties we will use all sort of tricks with σ matrices. So, to begin, let's review some of these tricks. From there, we will construct explicit representation matrices and build Lorentz-invariant combinations of objects that we might someday hope to add to a Lagrangian.

6.1.1 Tricks with σ Matrices

The Pauli σ matrices σ^a are defined to be

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (6.1)$$

They satisfy the handy relations

$$[\sigma^a, \sigma^b] = 2i\epsilon^{abc}\sigma^c, \quad (6.2)$$

$$\{\sigma^a, \sigma^b\} = 2\delta^{ab}. \quad (6.3)$$

The first of these relations implies that $\{\sigma^a/2\}$ form a two-dimensional representation of the Lie algebra of $SU(2)$. Taken together, Eqs. (6.2,6.3) also imply that

$$e^{-i\alpha^a\sigma^a/2} = \cos(\alpha/2) - i\sigma^a(\alpha^a/\alpha) \sin(\alpha/2) , \quad (6.4)$$

where $\alpha = \sqrt{\alpha^a\alpha^a}$. To derive this, expand the exponential and use Eqs. (6.2,6.3) together with $2AB = \{A, B\} + [A, B]$ for any operators A and B .

The Pauli matrices are clearly Hermitian, $\sigma^{a\dagger} = \sigma^a$. However, $(\sigma^a)^* = (\sigma^a)^t$ are not equal to σ^a when $a = 2$. To handle such conjugation operations, let us define another matrix

$$\epsilon = i\sigma^2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} . \quad (6.5)$$

This matrix is useful because

$$(\sigma^a)^* \epsilon = -\epsilon \sigma^a, \quad (\sigma^a)^t \epsilon = -\epsilon \sigma^a . \quad (6.6)$$

We call this the ϵ *trick*. The ϵ *trick* also works if we replace ϵ by $\bar{\epsilon}$ defined to be

$$\bar{\epsilon} = -\epsilon = \epsilon^t . \quad (6.7)$$

Note as well that $\epsilon^2 = -\mathbb{I} = \bar{\epsilon}^2$ and $\epsilon\bar{\epsilon} = \mathbb{I}$

Consider a general 2×2 matrix given by

$$M(\alpha^a) = e^{-i\alpha^a\sigma^a/2} , \quad (6.8)$$

for some parameters α^a . When the parameters are real, this matrix is clearly unitary. However, when the α^a are complex, M is no longer unitary in general. Despite the loss of unitarity, we can still find the inverse of M in a clever way by using the ϵ *trick* and Eq. (6.4). Together, they imply that

$$M^t(\alpha^a)\epsilon = \epsilon M(-\alpha^a) = \epsilon M^{-1}(\alpha^a) . \quad (6.9)$$

This result will be useful soon.

6.1.2 Building the $(1/2, 0)$ Representation

Recall that we had for a field $\phi_C(x)$ transforming under a rep r of the Lorentz group

$$U(\Lambda)^\dagger \phi_C(x) U(\Lambda) = [M(\Lambda)]_C^D \phi_D(\Lambda^{-1}x) . \quad (6.10)$$

The transformation matrix M can be written as an exponential,

$$M(\Lambda) = \exp\left(-\frac{i}{2}\omega_{\mu\nu}J_r^{\mu\nu}\right) = M(\omega) , \quad (6.11)$$

where the representation matrices $J_r^{\mu\nu}$ satisfy the commutation relations of the Lorentz group. Note as well that specifying Λ is equivalent to specifying the parameters $\omega_{\mu\nu}$. In

particular, given $\omega_{\mu\nu}$, the corresponding matrix is $\Lambda = \exp(-i\omega_{\mu\nu}J_4^{\mu\nu}/2)$ where we defined the generators $J_4^{\mu\nu}$ of the 4-vector rep in notes-05.

We also defined

$$J^i = \frac{1}{2}\epsilon^{ijk}J^{jk}, \quad K^i = J^{0i}. \quad (6.12)$$

In terms of them, we formed

$$A^i = \frac{1}{2}(J^i + iK^i), \quad B^i = \frac{1}{2}(J^i - iK^i). \quad (6.13)$$

The commutators of the A s and B s formed two independent sets of $SU(2)$ Lie algebras. The representations are therefore (j_A, j_B) , where j_A and j_B refer to the dimension of each component.

To form the $(1/2, 0)$ rep, we use

$$A^i = \sigma^i/2, \quad B^i = 0. \quad (6.14)$$

In terms of these, we can solve for the matrices J^i and K^i in this rep:

$$J^i = A^i + B^i = \sigma^i/2, \quad (6.15)$$

$$K^i = -i(A^i - B^i) = -i\sigma^i/2. \quad (6.16)$$

It follows that

$$M(\omega) = \exp\left(-\frac{i}{2}\omega_{\mu\nu}J_{(1/2,0)}^{\mu\nu}\right) \quad (6.17)$$

$$= e^{-i(\theta^a - i\beta^a)\sigma^a/2} \quad (6.18)$$

$$= e^{-i\alpha^a\sigma^a/2} = M(\alpha^a), \quad (6.19)$$

where we have defined

$$\theta^a = \frac{1}{2}\epsilon^{abc}\omega_{bc}, \quad \beta^a = \omega_{0a}. \quad (6.20)$$

The 2×2 matrix $M(\omega)$ looks just like a regular $SU(2)$ transformation, but now with a set of three *complex* parameters $\alpha^a = (\theta^a - i\beta^a)$. The real part of α^a corresponds to a rotation about the a -th spatial axis, and the imaginary part corresponds to a boost in the a -th spatial direction. Note that specifying Λ , $\omega_{\mu\nu}$, or α^a all provide equivalent ways to define a Lorentz transformation and we will use them interchangeably.

The Lorentz transformation of a $(1/2, 0)$ spinor ψ_α is therefore

$$U^\dagger(\alpha^a)\psi_\alpha(x)U(\alpha^a) = [M(\alpha^a)]_\alpha^\beta\psi_\beta(\Lambda^{-1}x), \quad (6.21)$$

where $\alpha, \beta = 1, 2$. The use of Greek indices to label the components of the spinor ψ_α is traditional but unfortunate – make sure you don't confuse them with 4-vector indices.

Given the form of $M(\alpha^a)$, we can build a Lorentz-invariant bilinear operator. Using our ϵ trick, we have

$$M^t(\alpha^a)\epsilon M(\alpha^a) = \epsilon . \quad (6.22)$$

This implies that given any two $(1/2, 0)$ spinors ψ and χ , the combination $[\chi^t\epsilon\psi]$ is Lorentz-invariant. Putting in indices,

$$[\chi^t\epsilon\psi] = \chi_\beta\epsilon^{\beta\alpha}\psi_\alpha = -(\epsilon^{\alpha\beta}\chi_\beta)\psi_\alpha . \quad (6.23)$$

For this reason, it is standard to define a spinor with a raised index,

$$\chi^\alpha := \epsilon^{\alpha\beta}\chi_\beta . \quad (6.24)$$

In terms of this, the Lorentz-invariant bilinear is written as

$$\chi\psi := \chi^\alpha\psi_\alpha . \quad (6.25)$$

We would also like to be able to lower the spinor index. Using $\epsilon\bar{\epsilon} = 1$, it follows that we can do this with $\bar{\epsilon}_{\alpha\beta}$:

$$\psi_\alpha = \bar{\epsilon}_{\alpha\beta}\psi^\beta . \quad (6.26)$$

The bar on $\bar{\epsilon}$ is usually not written explicitly. Instead, the standard notation has $\epsilon^{\alpha\beta}$ antisymmetric with $\epsilon^{12} = +1$, and $\epsilon_{\alpha\beta}$ also antisymmetric with $\epsilon_{12} = -1$. Thus, $\epsilon^{\alpha\lambda}\epsilon_{\lambda\beta} = \delta^\alpha_\beta$.

6.1.3 Building the $(0,1/2)$ and Representation and More

Having constructed the $(1/2, 0)$ representation, the $(0, 1/2)$ rep is really easy. We have

$$A^i = 0, \quad B^i = \sigma^i/2 , \quad (6.27)$$

which implies

$$J^i = \sigma^i/2 \quad (6.28)$$

$$K^i = i\sigma^i/2 . \quad (6.29)$$

A general finite element is therefore

$$\bar{M}(\alpha^a) = e^{-i(\theta^a + i\beta^a)\sigma^a/2} \quad (6.30)$$

$$= e^{-i(\alpha^a)^*\sigma^a/2} , \quad (6.31)$$

where θ^a and β^a are related to $\omega_{\mu\nu}$ in exactly the same way as before.

Given the similarity of this form to the $(1/2, 0)$ rep, we will use a peculiar but ultimately useful notation for the indices of a $(0,1/2)$ spinor $\bar{\psi}(x)$ (where the bar on the field is part of its name, not some sort of conjugation operation):

$$U^\dagger(\alpha^a)\bar{\psi}(x)U(\alpha^a) = [\bar{M}(\alpha^a)]^{\dot{\alpha}}_{\dot{\beta}}\bar{\psi}^{\dot{\beta}}(\Lambda^{-1}x) , \quad (6.32)$$

where the indices $\dot{\alpha}, \dot{\beta} = 1, 2$. With this transformation law, we can immediately form a Lorentz-invariant bilinear operator from a pair of $(0, 1/2)$ spinors $\bar{\chi}$ and $\bar{\psi}$ using our ϵ *trick*:

$$\bar{\chi}\bar{\psi} := \bar{\chi}_{\dot{\alpha}}\bar{\psi}^{\dot{\alpha}} = (\epsilon_{\dot{\alpha}\dot{\beta}}\bar{\chi}^{\dot{\beta}})\bar{\psi}^{\dot{\alpha}}, \quad (6.33)$$

where $\epsilon_{\dot{\alpha}\dot{\beta}}$ corresponds to the matrix $\bar{\epsilon}$. Similarly, we can raise indices using $\epsilon^{\dot{\alpha}\dot{\beta}}$,

$$\bar{\chi}^{\dot{\alpha}} = \epsilon^{\dot{\alpha}\dot{\beta}}\bar{\chi}_{\dot{\beta}}. \quad (6.34)$$

The components of $\epsilon^{\dot{\alpha}\dot{\beta}}$ are equal to those of $\epsilon^{\alpha\beta}$, and the same for $\epsilon_{\dot{\alpha}\dot{\beta}}$ and $\epsilon_{\alpha\beta}$.

The notation we are using looks funny, but there is a good reason for it. Consider the transformation property of the $(1/2, 0)$ spinor ψ with a raised index,

$$\psi^{\alpha} \rightarrow \epsilon^{\alpha\lambda}[M(\alpha^a)]_{\lambda}^{\beta}\psi_{\beta} \quad (6.35)$$

$$= \epsilon^{\alpha\lambda}[M(\alpha^a)]_{\lambda}^{\beta}\epsilon_{\beta\kappa}\psi^{\kappa} \quad (6.36)$$

$$= \left[e^{+i\alpha^a(\sigma^a)^t/2} \right]_{\kappa}^{\alpha} \psi^{\kappa}, \quad (6.37)$$

where we have arranged the index structure of the last matrix to make it consistent. Comparing Eq. (6.37) to the transformation of Eq. (6.31), we see that $(\psi^{\alpha})^*$ transforms in exactly the same way under Lorentz as a $(0, 1/2)$ spinor!

Given a $(1/2, 0)$ spinor ψ_{α} , we can therefore construct a $(0, 1/2)$ spinor $\bar{\psi}^{\dot{\alpha}}$ by

$$\bar{\psi}^{\dot{\alpha}} := \epsilon^{\dot{\alpha}\dot{\beta}}(\psi^*)_{\dot{\beta}}, \quad (6.38)$$

where we have written $(\psi^*)_{\dot{\beta}} = (\psi_{\beta})^*$. Similarly, given a $(0, 1/2)$ spinor $\bar{\chi}^{\dot{\alpha}}$, we can form a $(1/2, 0)$ spinor through

$$\chi_{\alpha} := \epsilon_{\alpha\beta}(\bar{\chi}^*)^{\beta}. \quad (6.39)$$

Thanks to these handy relations, we only really ever need to deal with $(1/2, 0)$ spinors.

For our next trick, let us try to connect the $(1/2, 1/2)$ rep with the 4-vector rep. To do so, let us define the set of four 2×2 matrices σ^{μ} ($\mu = 0, 1, 2, 3$) by

$$\sigma_{\alpha\dot{\alpha}}^{\mu} = (\mathbb{I}, \vec{\sigma})_{\alpha\dot{\alpha}}. \quad (6.40)$$

With this, we can form the object

$$\psi\sigma^{\mu}\bar{\chi} := \psi^{\alpha}\sigma_{\alpha\dot{\alpha}}^{\mu}\bar{\chi}^{\dot{\alpha}}. \quad (6.41)$$

Under Lorentz, we have

$$\psi^{\alpha}\sigma_{\alpha\dot{\alpha}}^{\mu}\bar{\chi}^{\dot{\alpha}} \rightarrow \epsilon^{\alpha\lambda}(M_{\lambda}^{\beta}\psi_{\beta})\sigma_{\alpha\dot{\alpha}}^{\mu}(\bar{M}^{\dot{\alpha}}_{\dot{\beta}}\bar{\chi}^{\dot{\beta}}) \quad (6.42)$$

$$= -\psi_{\beta}[M^t\epsilon\sigma^{\mu}\bar{M}]^{\beta}_{\dot{\beta}}\bar{\chi}^{\dot{\beta}} \quad (6.43)$$

In the homework you will show that

$$[M^t \epsilon \sigma^\mu \overline{M}]^{\beta}_{\dot{\beta}} = \epsilon^{\beta\lambda} \Lambda^\mu_{\nu} \sigma^\nu_{\lambda\dot{\beta}} , \quad (6.44)$$

where Λ is the corresponding 4-vector transformation. It follows that under Lorentz

$$\psi \sigma^\mu \bar{\chi} \rightarrow \Lambda^\mu_{\nu} \psi \sigma^\nu \bar{\chi} , \quad (6.45)$$

which transforms like a 4-vector. This justifies our notation for the upper index on σ^μ , and it also shows how the 4-vector rep emerges from the $(1/2, 1/2)$ rep.

Let us also define

$$(\bar{\sigma}^\mu)^{\dot{\alpha}\alpha} = (\mathbb{I}, -\vec{\sigma})^{\dot{\alpha}\alpha} . \quad (6.46)$$

Following the same steps as before, one can show that for any $(1/2, 0)$ and $(0, 1/2)$ spinors ψ and $\bar{\chi}$,

$$\bar{\chi} \bar{\sigma}^\nu \psi := \bar{\chi}_{\dot{\alpha}} (\bar{\sigma}^\mu)^{\dot{\alpha}\alpha} \psi_\alpha \quad (6.47)$$

$$\rightarrow \Lambda^\mu_{\nu} \bar{\chi} \bar{\sigma}^\nu \psi . \quad (6.48)$$

Numerically, one also has

$$(\bar{\sigma}^\mu)^{\dot{\alpha}\alpha} = \epsilon^{\dot{\alpha}\beta} \epsilon^{\alpha\beta} \sigma^\mu_{\beta\dot{\beta}} . \quad (6.49)$$

This relation means that the spinor indices on $\bar{\sigma}^\mu$ are consistent with raising and lowering with our good friend ϵ .

The σ^μ matrices also satisfy two very useful relations. The first follows from the tracelessness of the Pauli matrices, and reads

$$tr(\bar{\sigma}^\mu \sigma^\nu) = (\bar{\sigma}^\mu)^{\dot{\alpha}\alpha} \sigma^\nu_{\alpha\dot{\alpha}} = 2\eta^{\mu\nu} . \quad (6.50)$$

The second relation is

$$\sigma^\mu_{\alpha\dot{\alpha}} (\sigma_\mu)_{\beta\dot{\beta}} = -2\epsilon_{\alpha\beta} \epsilon_{\dot{\alpha}\dot{\beta}} . \quad (6.51)$$

6.1.4 Spinor Lagrangians

With all that spinor technology out of the way, we are now able to put together Lorentz-invariant Lagrangians for spinor fields. In doing so, however, there are two additional conditions. First, to describe a physical system, the action must be real. Since a spinor ψ is an inherently complex object, we must therefore have $\bar{\psi} = \epsilon\psi^*$ in our theory as well. Second, when we quantize later on we will find that spinors describe fermions. It turns out that for the quantum theory to connect in a reasonable way to a classical theory, the spinors must anticommute with each other (even in the classical theory). In particular,

$$\psi_\alpha \chi_\beta = -\chi_\beta \psi_\alpha, \quad \psi_\alpha \bar{\chi}_{\dot{\beta}} = -\bar{\chi}_{\dot{\beta}} \psi_\alpha . \quad (6.52)$$

In fancy math language, spinors are said to be *Grassmann variables*.

For future use, it will be useful to have an explicit convention for the complex conjugation of multiple classical fields, whether they be bosonic (and commuting) or fermionic (and anticommuting). For either field type, we define for a single field

$$A^\dagger := A^* , \quad (6.53)$$

as well as

$$(A_1 A_2 \dots A_n)^* := (A_1 A_2 \dots A_n)^\dagger := A_n^\dagger \dots A_2^\dagger A_1^\dagger . \quad (6.54)$$

Note that we have reversed the order with no additional signs, even for the fermion case. This convention is useful because it will match smoothly with the operation of Hermitian conjugation in the quantum theory, where we promote the fields to operators on a Hilbert space. Note as well that

$$(\chi\xi)^\dagger = (\epsilon^{\alpha\beta}\chi_\beta\xi_\alpha)^\dagger = \epsilon^{\dot{\alpha}\dot{\beta}}\bar{\xi}_{\dot{\alpha}}\bar{\chi}_{\dot{\beta}} = +\bar{\xi}\bar{\chi} . \quad (6.55)$$

At the very least, a sensible physical theory requires a kinetic term involving some spacetime derivatives. It turns out that the right form for a spinor is

$$\mathcal{L} \supset \frac{1}{2}\psi i\sigma^\mu \partial_\mu \bar{\psi} + \frac{1}{2}\bar{\psi} i\bar{\sigma}^\mu \partial_\mu \psi \quad (6.56)$$

$$= \bar{\psi} i\bar{\sigma}^\mu \partial_\mu \psi = \psi i\sigma^\mu \partial_\mu \bar{\psi} \quad (6.57)$$

where you will verify the reality of the first line and the equalities in the second line in the homework.

We can also add a bilinear mass term for the spinor. If we only have a single spinor ψ (and its conjugate $\bar{\psi}$), the only Lorentz-invariant possibility is

$$\mathcal{L} \supset -\frac{1}{2}m\psi\psi - \frac{1}{2}m^*\bar{\psi}\bar{\psi} . \quad (6.58)$$

You might worry that these terms both vanish since ψ is anticommuting, but they do not. Note that

$$\chi\psi = \chi^\alpha\psi_\alpha = \epsilon^{\alpha\beta}\chi_\beta\psi_\alpha = -\epsilon^{\alpha\beta}\psi_\alpha\chi_\beta = \epsilon^{\beta\alpha}\psi_\alpha\chi_\beta = \psi^\beta\chi_\beta = \psi\chi , \quad (6.59)$$

where we see that the anticommutation of the spinors is cancelled by the antisymmetry of ϵ .

6.2 Dirac Fermions

Having spent all that time on two-component $(1/2, 0)$ and $(0, 1/2)$ spinors, we turn next to study four-component objects in the $(1/2, 0) \oplus (0, 1/2)$ representation. While the rep is reducible under Lorentz, it is irreducible if we also impose parity which exchanges j_A and j_B in (j_A, j_B) . Parity turns out to be a good symmetry of electromagnetism, and therefore we would like to build it into our fields. This is why we'll use four-component Dirac fermions to describe electrons in QED.

6.2.1 Dirac Spinors

Consider a theory containing two $(1/2, 0)$ spinors ξ and χ together with their conjugates. We will assume the theory has a global symmetry under the phase transformations

$$\xi(x) \rightarrow e^{-i\varphi}\psi(x), \quad \chi(x) \rightarrow e^{i\varphi}\chi(x) . \quad (6.60)$$

The most general Lagrangian for this theory at bilinear order is

$$\mathcal{L} = \bar{\xi}i\bar{\sigma}^\mu\partial_\mu\xi + \bar{\chi}i\bar{\sigma}^\mu\partial_\mu\chi - m(\xi\chi + \bar{\xi}\bar{\chi}) . \quad (6.61)$$

The global symmetry allows the mixed $\chi\xi$ quadratic term, but it forbids $\chi\chi$ and $\xi\xi$.

When we quantize the theory, we will interpret m as the mass of a particle. However, since two fields are involved, it is not obvious how to relate the mass term to a specific particle. Using a four-component spinor containing both two-component spinors allows us to dodge this issue for the time being. We define the four-component *Dirac spinor* Ψ by

$$\Psi = \begin{pmatrix} \xi_\alpha \\ \bar{\chi}^{\dot{\alpha}} \end{pmatrix} . \quad (6.62)$$

The conjugate of Ψ is thus

$$\Psi^\dagger = (\bar{\xi}_{\dot{\alpha}}, \chi^\alpha) . \quad (6.63)$$

To go along with Ψ , we also generalize the σ^μ matrices to the 4×4 Dirac matrices γ^μ ,

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix} , \quad (6.64)$$

where each of the matrix elements is itself a 2×2 matrix. Finally, let us define the barred conjugate $\bar{\Psi}$ to be

$$\bar{\Psi} = \Psi^\dagger\gamma^0 = (\chi^\alpha, \bar{\xi}_{\dot{\alpha}}) . \quad (6.65)$$

Note that here the bar denotes a conjugation operation, and is not part of the name of the Dirac spinor.

With these definitions in hand, we can rewrite the Lagrangian of Eq. (6.61) in a more compact form using Ψ . The result is

$$\mathcal{L} = \bar{\Psi}i\gamma^\mu\partial_\mu\Psi - m\bar{\Psi}\Psi . \quad (6.66)$$

The mass term looks much nicer now.

6.2.2 Fun with γ Matrices

The γ matrices satisfy all kinds of useful relations that we will make use of when we compute scattering amplitudes in QED. The most important one is

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} . \quad (6.67)$$

It is easy to verify this relation using the definition of Eq. (6.64) and the properties of σ matrices. More generally, the γ matrices can be defined in any number of spacetime dimensions to be the minimal solutions of Eq. (6.67). In four dimensions, all the solutions can be shown to be equivalent to Eq. (6.64) up to a unitary change of basis for the four-component space.

Many of the other useful relations involve traces of γ matrices. Recall that the trace of any matrix M is defined to be

$$\text{tr}(M) = \sum_i M_{ii} = M_{11} + M_{22} + \dots \quad (6.68)$$

Traces are linear,

$$\text{tr}(aM) = a \text{tr}(M) , \quad (6.69)$$

and they are cyclic,

$$\text{tr}(AB) = \text{tr}(BA), \quad \text{tr}(ABC) = \text{tr}(BCA) = \text{tr}(CAB), \quad \text{etc.} \quad (6.70)$$

From the explicit form of Eq. (6.64), we see that

$$\text{tr}(\gamma^\mu) = 0 . \quad (6.71)$$

This can be generalized, and it turns out that the trace of any odd number of γ^μ matrices vanishes. On the other hand, for even numbers we have

$$\text{tr}(\gamma^\mu \gamma^\nu) = 4\eta^{\mu\nu} \quad (6.72)$$

$$\text{tr}(\gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma) = 4(\eta^{\mu\nu} \eta^{\rho\sigma} - \eta^{\mu\rho} \eta^{\nu\sigma} + \eta^{\mu\sigma} \eta^{\nu\rho}) , \quad (6.73)$$

which can both be derived using the cyclicity of the trace. An additional useful relation is

$$p_\mu p_\nu \gamma^\mu \gamma^\nu = \frac{1}{2} p_\mu p_\nu (\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu) = 2\eta^{\mu\nu} p_\mu p_\nu = p^2 . \quad (6.74)$$

Finally, we have

$$\gamma^\mu \gamma_\mu = \eta_{\mu\nu} \gamma^\mu \gamma^\nu = \eta^{\mu\nu} \eta_{\mu\nu} = 4 , \quad (6.75)$$

at least in four spacetime dimensions.

Let us also define the matrix $\gamma^5 = \gamma_5$ to be

$$\gamma^5 = i\gamma^0 \gamma^1 \gamma^2 \gamma^3 . \quad (6.76)$$

This matrix anticommutes with all the γ^μ ,

$$\{\gamma^5, \gamma^\mu\} = 0 , \quad (6.77)$$

and satisfies

$$(\gamma^5)^2 = 1 . \quad (6.78)$$

Using the basis of Eq. (6.64), the explicit form of γ^5 is

$$\gamma^5 = \begin{pmatrix} -\mathbb{I} & 0 \\ 0 & \mathbb{I} \end{pmatrix} . \quad (6.79)$$

It is also conventional to define the projectors P_L and P_R to be

$$P_{L,R} = \frac{1}{2}(1 \mp \gamma^5) . \quad (6.80)$$

Applying this to a Dirac spinor, we get

$$P_L \Psi = \Psi_L = \begin{pmatrix} \xi_\alpha \\ 0 \end{pmatrix}, \quad P_R \Psi = \Psi_R = \begin{pmatrix} 0 \\ \bar{\chi}^{\dot{\alpha}} \end{pmatrix} . \quad (6.81)$$

Note that

$$\Psi = \Psi_L + \Psi_R , \quad (6.82)$$

and

$$\bar{\Psi} i \gamma^\mu \partial_\mu \Psi = \bar{\Psi}_L i \gamma^\mu \partial_\mu \Psi_L + \bar{\Psi}_R i \gamma^\mu \partial_\mu \Psi_R , \quad (6.83)$$

as well as

$$\bar{\Psi} \Psi = \bar{\Psi}_L \Psi_R + \bar{\Psi}_R \Psi_L . \quad (6.84)$$

Neither result should be surprising given what you know about two-component spinors.

We will also sometimes evaluate traces involving γ^5 . Clearly, we have

$$\text{tr}(\gamma^5) = 0 . \quad (6.85)$$

Since the trace of an odd number of γ matrices vanishes, we also have

$$\text{tr}(\gamma^5 \times (\text{odd number of } \gamma^\mu\text{s})) = 0 . \quad (6.86)$$

Cyclicity also implies

$$\text{tr}(\gamma^5 \gamma^\mu \gamma^\nu) = 0 . \quad (6.87)$$

However, the trace with four γ matrices is non-zero,

$$\text{tr}(\gamma^5 \gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma) = -4i \epsilon^{\mu\nu\rho\sigma} , \quad (6.88)$$

where $\epsilon^{\mu\nu\rho\sigma}$ is completely antisymmetric with $\epsilon^{0123} = +1$.

Chapter 7

Classical and Quantum Fermions

With the spinor technology in place, we can get to work building a quantum field theory with them. We will find that the only consistent way to do so is to have the spinors represent fermions.

7.1 Classical Fermions

As a warmup, let's start off by studying the classical equations of motion.

7.1.1 Classical Weyl

The simplest Lagrangian we can have is

$$\mathcal{L} = \bar{\psi} i \bar{\sigma}^\mu \partial_\mu \psi , \quad (7.1)$$

where $\psi(x)$ transforms in the $(1/2, 0)$ representation of Lorentz. In finding the equations of motion, we should treat ψ and $\bar{\psi}$ as independent variables. This gives

$$\frac{\delta S}{\delta \bar{\psi}_\alpha(x)} = i(\bar{\sigma}^\mu)^{\dot{\alpha}\alpha} \partial_\mu \psi_\alpha = i(\bar{\sigma} \cdot \partial \psi)^{\dot{\alpha}} \quad (7.2)$$

$$\frac{\delta S}{\delta \psi_\alpha(x)} = -i \partial_\mu \bar{\psi}_\alpha (\bar{\sigma}^\mu)^{\dot{\alpha}\alpha} = -i(\bar{\psi} \bar{\sigma} \cdot \overleftarrow{\partial})^\alpha . \quad (7.3)$$

Note that the indices match up nicely. We will suppress them below, but it's a good exercise to check that they all work out.

Let us try to find a plane wave solution of the form

$$\psi_\alpha(x) = u_\alpha(p) e^{-ip \cdot x} . \quad (7.4)$$

Putting this into the equations of motion, we find

$$(\bar{\sigma} \cdot p) u(p) = 0 . \quad (7.5)$$

Multiplying by $(\sigma \cdot p)$ and using $(\sigma \cdot p)(\bar{\sigma} \cdot p) = p^2$, we find

$$p^2 = 0 . \quad (7.6)$$

Thus $p^0 = \pm|\vec{p}|$, the dispersion relation of a massless particle. Consider the special case of momentum along the z axis: $p^\mu = k^\mu = (k, 0, 0, k)$. We get

$$0 = (\bar{\sigma} \cdot k)u(k) = k \left[\mathbb{I} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right] \begin{pmatrix} u_1(k) \\ u_2(k) \end{pmatrix} = k \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u_1(k) \\ u_2(k) \end{pmatrix} , \quad (7.7)$$

which implies $u_1(k) = 0$. An explicit solution is therefore $\psi(x) = u(k)e^{-ik \cdot x}$ with $k^2 = 0$ and

$$u(k) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} . \quad (7.8)$$

This coincides with what we would expect for a massless particle with spin anti-parallel to the direction of motion, which is said to be *left-handed*. Repeating these steps for $\bar{\psi}$ yields

$$\bar{\psi}^{\dot{\alpha}} = \bar{u}^{\dot{\alpha}}(k)e^{ik \cdot x} , \quad (7.9)$$

with

$$\bar{u}(k) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} . \quad (7.10)$$

This corresponds to what we would expect for a *right-handed* particle, with spin parallel to the direction of motion. Note as well that with our choice of normalization,

$$\bar{u}^{\dot{\alpha}} = \epsilon^{\dot{\alpha}\dot{\beta}}(u_\beta)^* . \quad (7.11)$$

This coincides with the expected relation between ψ and $\bar{\psi}$. It is straightforward to generalize this solution to arbitrary p (with $p^2 = m^2$).

Recall that in **notes-06** we showed that a massless particle with positive energy has a single helicity state while its antiparticle has the opposite helicity. This agrees with the classical solutions we've found for $k^0 > 0$. The spinor ψ is a two-component complex object, and it would naively have four real degrees of freedom. However, applying the equations of motion restricts the components of ψ beyond just the dispersion relation for k^0 . This yields two independent solutions that we will identify with a left-handed particle and a right-handed antiparticle when we quantize.¹

We could do much more with the classical Weyl fermion. However, since we are mainly interested in the electron of QED, which is described by a Dirac fermion consisting of a pair of Weyl fermions, we will move on to that case. For an exhaustive discussion of Weyl fermions, see Ref. [17].

¹The choice of what we label as particle or antiparticle is arbitrary.

7.1.2 Classical Dirac

The starting point for the theory of a (classical) Dirac fermion is the Lagrangian

$$\mathcal{L} = \bar{\Psi} i \gamma^\mu \partial_\mu \Psi - m \bar{\Psi} \Psi . \quad (7.12)$$

To derive the equations of motion, it will be convenient to define an index structure for these four-spinors. We will use

$$\Psi = \Psi_a, \quad \bar{\Psi} = \bar{\Psi}^a, \quad \gamma^\mu = (\gamma^\mu)_a{}^b , \quad (7.13)$$

where $a = 1, 2, 3, 4$. The equations of motion for Ψ and $\bar{\Psi}$ are

$$\frac{\delta S}{\delta \bar{\Psi}^a(x)} = 0 = (i \gamma^\mu \partial_\mu - m)_a{}^b \Psi_b(x) , \quad (7.14)$$

$$\frac{\delta S}{\delta \Psi_a(x)} = 0 = \bar{\Psi}^a(x) (-i \gamma^\mu \overleftarrow{\partial}_\mu - m)_a{}^b . \quad (7.15)$$

The first equation is called the *Dirac equation* and the second is just its conjugate.

As before, let's look for plane wave solutions to these equations,

$$\Psi_a(x) = a(p) u_a(p) e^{-ip \cdot x} , \quad (7.16)$$

where $a(p)$ is a Grassmannian normalization factor and $u_a(p)$ is a four-component column vector of regular (non-Grassmann) numbers, and we will restrict $p^0 > 0$. The Dirac equation then becomes

$$(\gamma^\mu p_\mu - m)u(p) = (\not{p} - m)u(p) = 0 . \quad (7.17)$$

In this expression we have introduced the Feynman slash notation:

$$\not{v} = \gamma^\mu v_\mu = \gamma \cdot v , \quad (7.18)$$

for any four-vector v^μ . Multiplying by $(\not{p} + m)$, we find that

$$(p^2 - m^2)u(k) = 0 , \quad (7.19)$$

which implies $p^0 = \pm \sqrt{\vec{p}^2 + m^2}$, as one would expect for a particle of mass m . Going back to Eq. (7.17), we can rewrite the result in matrix form as

$$0 = \begin{pmatrix} -m & \sigma \cdot p \\ \bar{\sigma} \cdot p & -m \end{pmatrix} \begin{pmatrix} u_L \\ u_R \end{pmatrix} , \quad (7.20)$$

where we have split u into two two-component pieces. Expanding gives

$$u_R(p) = \frac{\bar{\sigma} \cdot p}{m} u_L(p) , \quad (7.21)$$

which relates the components of $u_L(p)$ and $u_R(p)$ in a non-trivial way. If we plug this into the second equation, we get

$$u_L(p) = \frac{1}{m^2}(\sigma \cdot p)(\bar{\sigma} \cdot p)u_L(p) = \frac{p^2}{m^2}u_L(p) , \quad (7.22)$$

which does not relate the components any further but forces $p^2 = m^2$.

We would like to find a simple basis for these plane wave solutions. It would also be nice if it were somewhat symmetric between u_L and u_R . A convenient way to do this turns out to be [1]

$$u(p, s) = \begin{pmatrix} \sqrt{\sigma \cdot p} \xi(s) \\ \sqrt{\bar{\sigma} \cdot p} \xi(s) \end{pmatrix} , \quad (7.23)$$

where the square root of the matrix is defined as a formal power series, and $s = 1, 2$ label the two independent basis solutions

$$\xi(1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \quad \xi(2) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} . \quad (7.24)$$

This form looks funny, but it is consistent with Eq. (7.21); just plug Eq. (7.23) into Eq. (7.21) and use $(\sigma \cdot p)(\bar{\sigma} \cdot p) = p^2$:

$$\frac{\bar{\sigma} \cdot p}{m} u_L(p, s) = \frac{1}{m} \sqrt{\bar{\sigma} \cdot p} \sqrt{(\bar{\sigma} \cdot p)(\sigma \cdot p)} \xi(s) = \sqrt{\bar{\sigma} \cdot p} \xi(s) = u_R(p, s) . \quad (7.25)$$

Everything works out nicely!

These two solutions have some convenient properties. First, consider them in the rest frame, $p = (m, \vec{0})$, where they reduce to

$$u(p, s) = \sqrt{m} \begin{pmatrix} \xi(s) \\ \xi(s) \end{pmatrix} , \quad (7.26)$$

corresponding to spin up or spin down. More generally, in any frame the solutions are orthogonal,

$$\bar{u}(p, s)u(p, r) = u^\dagger(p, s)\gamma^0 u(p, r) = 2m \delta^{rs} . \quad (7.27)$$

Note that the normalization is Lorentz-invariant (just like $\bar{\Psi}\Psi$). The completeness of these solutions implies

$$\sum_{s=1}^2 u_a(p, s)\bar{u}^b(p, s) = (\gamma \cdot p + m)_a^b . \quad (7.28)$$

We will use this a lot in our calculations.

There are also two other independent solutions for “negative energy”, with $\tilde{p}^0 = -\sqrt{\vec{p}^2 + m^2}$. It can be shown that these are equivalent to solutions with $p^0 = +\sqrt{\vec{p}^2 + m^2}$ but with the sign flipped in the exponential:

$$\Psi_a(x) = b^*(p)v_a(p)e^{ip \cdot x} , \quad (7.29)$$

where $p^0 > 0$ as before. The equations of motion imply

$$-(\gamma^\mu p_\mu + m)v(p) = 0 . \quad (7.30)$$

A pair of independent solutions are

$$v(p, s) = \begin{pmatrix} \sqrt{\sigma \cdot p} \eta(s) \\ -\sqrt{\bar{\sigma} \cdot p} \eta(s) \end{pmatrix} , \quad (7.31)$$

with $s = 1, 2$ and

$$\eta(1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \quad \eta(2) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} . \quad (7.32)$$

In the rest frame, this reduces to

$$v(p, s) = \sqrt{m} \begin{pmatrix} \eta(s) \\ -\eta(s) \end{pmatrix} . \quad (7.33)$$

Similarly to before, we also have completeness in the form

$$\sum_{s=1}^2 v_a(p, s) \bar{v}^b(p, s) = (\gamma \cdot p - m)_a^b , \quad (7.34)$$

and orthogonality

$$\bar{v}(p, s)v(p, r) = -2m \delta^{sr} . \quad (7.35)$$

For future calculations it will be useful to have a list of other spinor products:

$$\bar{u}(p, r)u(p, s) = 2m \delta^{rs} \quad (7.36)$$

$$u^\dagger(p, r)u(p, s) = 2p^0 \delta^{rs} \quad (7.37)$$

$$\bar{v}(p, s)v(p, r) = -2m \delta^{rs} \quad (7.38)$$

$$v^\dagger(p, s)v(p, r) = 2p^0 \delta^{rs} \quad (7.39)$$

$$\bar{v}(p, r)u(p, s) = 0 \quad (7.40)$$

$$v^\dagger(p, s)u(p, r) = -2\eta^\dagger(s)(\vec{p} \cdot \vec{\sigma})\xi(r) \quad (7.41)$$

$$v^\dagger(p, r)u(p', s) = 0, \quad \text{for } p' = (p^0, -\vec{p}) . \quad (7.42)$$

These are straightforward to prove by simply multiplying the explicit forms of $u(p, s)$ and $v(p, r)$. For example,

$$\bar{v}(p, r)u(p, s) = \eta^\dagger(r) (\sqrt{\bar{\sigma} \cdot p} \sqrt{\sigma \cdot p} - \sqrt{\sigma \cdot p} \sqrt{\bar{\sigma} \cdot p}) \xi(s) \quad (7.43)$$

$$= (\sqrt{p^2} - \sqrt{p^2}) \eta^\dagger(r) \xi(s) . \quad (7.44)$$

Pay special attention to the sign flip in Eq. (7.38) (but not Eq. (7.39)).

With these basis solutions in hand, we can superpose them to get the most general solution:

$$\Psi(x) = \sum_{s=1}^2 \int \widetilde{d}p \left[a(p, s) u(p, s) e^{-ip \cdot x} + b^*(p, s) v(p, s) e^{ip \cdot x} \right] \quad (7.45)$$

$$\bar{\Psi}(x) = \sum_{s=1}^2 \int \widetilde{d}p \left[a^*(p, s) \bar{u}(p, s) e^{ip \cdot x} + b(p, s) \bar{v}(p, s) e^{-ip \cdot x} \right] . \quad (7.46)$$

In this expression, the exponentials carry the spacetime dependence, the factors of $u(p, s)$ and $v(p, s)$ carry the the four-spinor structure, and the expansion coefficients $a(p, s)$ and $b(p, s)$ make the right side Grassmanian.² In all, $\Psi(x)$ is a linear combination of four independent solutions (for each fixed \vec{p}) even though it would seem to have eight independent real degrees of freedom. What has happened is that four of these have been eliminated by the algebraic structure of the equations of motion. These remaining four coincide with what we would expect for a particle of spin $s = 1/2$ and an antiparticle with the same spin.

7.2 Quantum Fermions

Now we quantize! We will focus on the Dirac case for now.

7.2.1 Hamiltonian

Before quantizing, let us derive the classical Hamiltonian formulation of the theory defined by the Lagrangian of Eq. (7.12). The canonical momenta conjugate to $\Psi_a(x)$ and $\Psi^{\dagger a}(x)$ are

$$\Pi^a(x) = \frac{\partial \mathcal{L}}{\partial(\partial_0 \Psi_a(x))} = (\bar{\Psi} i \gamma^0)^a = i \Psi^{\dagger a} , \quad (7.47)$$

$$\bar{\Pi}_a(x) = \frac{\partial \mathcal{L}}{\partial(\partial_0 \Psi^{\dagger a}(x))} = 0 . \quad (7.48)$$

These look a bit funny, but this is just the result of having a kinetic term that is linear in derivatives. What they imply is that we only need to worry about the canonical pair Ψ and $i\Psi^\dagger$ when we quantize later on.

The Hamiltonian of the theory is therefore

$$H = \int d^3x \mathcal{H} = \int d^3x \left(\Pi^a \partial_0 \Psi_a - \mathcal{L} \right) \quad (7.49)$$

$$= \int d^3x \left(-\bar{\Psi} i \gamma^k \partial_k \Psi + m \bar{\Psi} \Psi \right) , \quad (7.50)$$

where $k = 1, 2, 3$ is summed over. This looks funny as well, but it is also fine.

²Note that u and v are not Grassmann numbers, and thus their components commute.

7.2.2 Hamilton, Dirac, and Fermi Walk into a (h)Bar...

To quantize the theory, we elevate $\Psi_a(x)$ and $\Pi^a(x)$ to operators on a Hilbert space and impose equal-time anticommutation relations,

$$\{\Psi_a(t, \vec{x}), \Pi^b(t, \vec{x}')\} = i\delta_a^b \delta^{(3)}(\vec{x} - \vec{x}') . \quad (7.51)$$

We will argue why we should use anticommutators rather than commutators later on. Using our explicit expression for Π^a , Eq. (7.51) is equivalent to

$$\{\Psi_a(t, \vec{x}), \Psi^{\dagger b}(t, \vec{x}')\} = \delta_a^b \delta^{(3)}(\vec{x} - \vec{x}') . \quad (7.52)$$

Let us now *define* the operators $a(p, s)$ and $a^\dagger(p, s)$ (and their conjugates) via [2]

$$a(p, s) = \int d^3x e^{ip \cdot x} \bar{u}(p, s) \gamma^0 \Psi(x) \Big|_{t=0} , \quad (7.53)$$

$$a^\dagger(p, s) = \int d^3x e^{-ip \cdot x} \bar{\Psi}(x) \gamma^0 u(p, s) \Big|_{t=0} , \quad (7.54)$$

as well as $b(p, s)$ and $b^\dagger(p, s)$ by [2]

$$b(p, s) = \int d^3x e^{ip \cdot x} \bar{\Psi}(x) \gamma^0 v(p, s) \Big|_{t=0} , \quad (7.55)$$

$$b^\dagger(p, s) = \int d^3x e^{-ip \cdot x} \bar{v}(p, s) \gamma^0 \Psi(x) \Big|_{t=0} . \quad (7.56)$$

In all these expressions, we have implicitly set $p^0 = +\sqrt{\vec{p}^2 + m^2}$. With some work, one can show that if we were to apply these operations to the right-hand sides of Eqs. (7.45,7.46), valid for the classical theory, we would indeed extract the a and b mode functions.

With the definitions of Eqs. (7.53,7.54,7.55,7.56), the anticommutation relations of Eq. (7.51) imply very simple relations for the a 's and b 's :

$$\{a(k, s), a(p, r)\} = 0 = \{b(k, s), b(p, r)\} = \{a(k, s), b^\dagger(p, r)\} \quad (7.57)$$

$$\{a(k, s), a^\dagger(p, r)\} = (2\pi)^3 2p^0 \delta^{(3)}(\vec{k} - \vec{p}) \delta^{rs} = \{b(k, s), b^\dagger(p, r)\} \quad (7.58)$$

These should remind you of the complex scalar theory you encountered in **hw-01**.

Just like for scalar fields, we will define the vacuum of the theory $|0\rangle$ to be the unique state annihilated by every $a(k, s)$ and every $b(k, s)$ operator,

$$a(k, s)|0\rangle = 0 = b(k, s)|0\rangle . \quad (7.59)$$

All other states can be obtained by applying powers of $a^\dagger(k, s)$ and $b^\dagger(k, s)$ to the vacuum. This gives general states of the form

$$\begin{aligned} |k_1, s_1; \dots; k_N, s_N : p_1, r_1; \dots; p_M, r_M\rangle = \\ a^\dagger(k_1, s_1) \dots a^\dagger(k_N, s_N) b^\dagger(p_1, r_1) \dots b^\dagger(p_M, r_M) |0\rangle . \end{aligned} \quad (7.60)$$

Unlike for scalars, however, the anticommutation relations imply that

$$[a^\dagger(k, s)]^2 = 0 = [b^\dagger(p, r)]^2 . \quad (7.61)$$

This means we can only put $n = 0$ or 1 particles in a given state with momentum p and spin s . This coincides precisely with the defining property of fermions that no two of them can occupy the same quantum state. We also see that there are two species of particles, corresponding to a^\dagger and b^\dagger . We will interpret them as a particle and an antiparticle.

In terms of the a and b operators, we can invert the relations of Eqs. (7.53,7.54,7.55,7.56) and solve for Ψ and $\bar{\Psi}$ at $t = 0$:

$$\Psi(0, \vec{x}) = \sum_{s=1}^2 \int \widetilde{dk} \left[a(k, s)u(k, s)e^{i\vec{k}\cdot\vec{x}} + b^\dagger(k, s)v(k, s)e^{-i\vec{k}\cdot\vec{x}} \right] \quad (7.62)$$

$$\bar{\Psi}(0, \vec{x}) = \sum_{r=1}^2 \int \widetilde{dp} \left[a^\dagger(p, r)\bar{u}(p, r)e^{-i\vec{p}\cdot\vec{x}} + b(p, r)\bar{v}(p, r)e^{i\vec{p}\cdot\vec{x}} \right] \quad (7.63)$$

With these in hand, we can now write the Hamiltonian in terms of the a and b operators. The calculation is somewhat involved, but the result is

$$H = \sum_{s=1}^2 \int \widetilde{dk} k^0 \left[a^\dagger(k, s)a(k, s) - b(k, s)b^\dagger(k, s) \right] \quad (7.64)$$

$$= \sum_{s=1}^2 \int \widetilde{dk} k^0 \left[a^\dagger(k, s)a(k, s) + b^\dagger(k, s)b(k, s) \right] - (\text{constant}) . \quad (7.65)$$

As before, we drop the constant by adding a constant to the Lagrangian to cancel it off. This result also shows why we needed anticommutators instead of commutators. The relative sign in Eq. (7.64) is flipped in going to Eq. (7.65) by anticommuting the b 's. If they had been commutators instead, the minus sign would not have been eliminated. With such a minus sign, states of the form $b^\dagger(k, s)|0\rangle$ would contribute negatively to the energy by an amount $-k^0$. Since this would be physically disastrous – by adding more and more antifermions we could lower the energy arbitrarily – we need anticommutators instead.

To obtain expressions for the fields at arbitrary times, we implement time evolution with the Hamiltonian. It is not hard to check that

$$e^{iHt} a(k, s) e^{-iHt} = e^{-ik^0 t} a(k, s), \quad e^{iHt} b(k, s) e^{-iHt} = e^{ik^0 t} b(k, s) . \quad (7.66)$$

This implies that

$$\Psi(t, \vec{x}) = e^{iHt} \Psi(0, \vec{x}) e^{-iHt} = \sum_{s=1}^2 \int \widetilde{dk} \left[a(k, s)u(k, s)e^{-ik\cdot x} + b^\dagger(k, s)v(k, s)e^{ik\cdot x} \right] \quad (7.67)$$

$$\bar{\Psi}(t, \vec{x}) = e^{iHt} \bar{\Psi}(0, \vec{x}) e^{-iHt} = \sum_{s=1}^2 \int \widetilde{dk} \left[a^\dagger(k, s)\bar{u}(k, s)e^{ik\cdot x} + b(k, s)\bar{v}(k, s)e^{-ik\cdot x} \right] \quad (7.68)$$

It is also clear that the $(N+M)$ -particle states we constructed above are eigenstates of the Hamiltonian with total energy $E = (\sum_{i=1}^N k^0 + \sum_{j=1}^M p_j^0)$.

The identification of a^\dagger as creating a fermion and b^\dagger an antifermion can be illuminated by looking at the $U(1)$ global symmetry of the theory. The Lagrangian of Eq. (7.12) is symmetric under the transformations

$$\Psi(x) \rightarrow e^{-iq\varphi}\Psi(x), \quad \bar{\Psi}(x) \rightarrow e^{iq\varphi}\bar{\Psi}(x), \quad (7.69)$$

for any constant transformation parameter φ . The conserved current can be shown to be

$$j^\mu = q \bar{\Psi} \gamma^\mu \Psi, \quad (7.70)$$

and the corresponding conserved charge is

$$Q = \int d^3x j^0 = q \sum_s \int \widetilde{dk} [a^\dagger(k, s)a(k, s) - b^\dagger(k, s)b(k, s)]. \quad (7.71)$$

The total charge Q of a state is therefore just q times the number of fermions minus the number of antifermions. Thus, each fermion contributes q to the total charge Q and each antifermion contributes $-q$. Note that we can also interpret $j^\mu/q = (n_\Psi, \vec{j}_\Psi)$, where n_Ψ is the number density of fermions minus antifermions and \vec{j}_Ψ is the spatial number current.

7.2.3 Propagating

For fermions, we modify the definition of time ordering to [1]

$$T\{\Psi(x) \overset{(-)}{\Psi}(x')\} = \begin{cases} \Psi(x) \overset{(-)}{\Psi}(x') & ; t > t' \\ - \overset{(-)}{\Psi}(x') \Psi(x) & ; t < t' \end{cases}. \quad (7.72)$$

The additional minus sign reflects the fermionic nature of Ψ and $\bar{\Psi}$, and allows for a smooth limit with $t \rightarrow t'$.

With this definition, we can compute all the two-point functions using the mode expansion of Eq. (7.67). It is immediately obvious that

$$\langle 0|T\{\Psi_a(x)\Psi_b(x')\}|0\rangle = 0 = \langle 0|T\{\bar{\Psi}^a(x)\bar{\Psi}^b(x')\}|0\rangle. \quad (7.73)$$

On the other hand, the mixed two-point function is non-zero. We find

$$\langle 0|T\{\Psi_a(x)\bar{\Psi}^b(x')\}|0\rangle = \int \widetilde{dk} \left[\Theta(t-t')(\not{k} + m)e^{-ik(x-x')} \right. \quad (7.74)$$

$$\left. - \Theta(t'-t)(\not{k} - m)e^{ik(x-x')} \right]. \quad (7.75)$$

Here, we have $k^0 = \sqrt{\vec{k}^2 + m^2}$. We can rewrite it as an expression with $k^0 \in (-\infty, \infty)$ using the same contour integral tricks as the scalar:

$$\langle 0|T\{\Psi_a(x)\bar{\Psi}^b(x')\}|0\rangle = \int \frac{d^4k}{(2\pi)^4} \frac{i(\not{k} + m)_a^b}{k^2 - m^2 + i\epsilon} e^{-ik \cdot (x-x')} \quad (7.76)$$

$$:= [S_F(x-x')]_a^b, \quad (7.77)$$

where $S_F(x - x')$ is called the *Feynman propagator*.

For future reference, let us take the Fourier transform of the two-point function,

$$\int d^4x e^{ip \cdot x} \int d^4x' e^{ip' \cdot x'} \langle \Psi_a(x) \bar{\Psi}^b(x') \rangle = (2\pi)^4 \delta^{(4)}(p + p') \frac{i(\not{k} + m)_a^b}{p^2 - m^2 + i\epsilon} \quad (7.78)$$

$$= (2\pi)^4 \delta^{(4)}(p + p') \left[\tilde{S}_F(p) \right]_a^b. \quad (7.79)$$

Note that

$$(\not{k} - m)(\not{k} + m) = (k^2 - m^2 + i\epsilon). \quad (7.80)$$

For this reason, $\tilde{S}_F(p)$ is sometimes written as

$$\left[\tilde{S}_F(p) \right]_a^b = \left[\frac{i}{\not{k} - m} \right]_a^b, \quad (7.81)$$

where this notation is just a shorthand for the fact that the propagator is the inverse $(\not{k} - m)$.

The Feynman propagator also has a nice physical interpretation. For $t > t'$, the only contribution comes from $\langle 0|aa^\dagger|0\rangle$. This coincides with a fermion with charge $+q$ propagating forward in time from x' to x . On the other hand, the only contribution for $t < t'$ comes from $\langle 0|bb^\dagger|0\rangle$, corresponding to an antifermion with charge $-q$ propagating forward in time from x to x' . Both possibilities are described by the amplitude $S_F(x - x')$. Now, for $t > t'$ this amplitude contains a time propagation factor of $\exp(-ip^0|t - t'|)$. However, for $t < t'$ we have a factor of $\exp(+ip^0|t - t'|)$, which seems to have the wrong sign. The insight of Feynman and Stückelberg was that a fermion travelling forward in time from x' to x is indistinguishable from an antifermion travelling backwards in time from x' to x . In both cases a charge $+q$ begins at \vec{x}' and finishes at \vec{x} . This accounts for the opposite sign on the exponential factor for $t < t'$.

Chapter 8

Interacting Fermions

We turn next to study theories with both fermions and interactions.

8.1 Perturbation Theory

As a specific example, we will study the interacting field theory of a real scalar ϕ and a Dirac fermion Ψ given by

$$\mathcal{L} = \bar{\Psi}i\partial\not{\Psi} - m\bar{\Psi}\Psi + \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}M^2\phi^2 - y\phi\bar{\Psi}\Psi . \quad (8.1)$$

The last term is evidently the interaction piece, $\Delta\mathcal{H} = \Delta V = y\phi\bar{\Psi}\Psi$. It is called a *Yukawa interaction* (and y is called the Yukawa coupling) after H. Yukawa who first suggested this form to describe the strong interaction between two nucleons (fermions) and a pion (scalar).

8.1.1 The Master Formula

We already know all the energy eigenstates of the theory in the limit of $y \rightarrow 0$. They are just free particle states containing some number of scalars, fermions, and antifermions, all with definite momentum (and spin for the fermions). However, as soon as we turn on the coupling, life becomes much more complicated. Even so, we can still study the interacting theory as an expansion in the coupling around the free theory. For this, we make two assumptions:

1. There exists a unique vacuum $|\Omega\rangle$ of the full theory with $p^\mu|\Omega\rangle = 0$.
2. For each field in the Lagrangian, there exists a set of distinct one-particle momentum eigenstates (possibly with several spin/helicity sub-states). Such states correspond to isolated poles in the two-point functions of the elementary fields.

Note that these are both assumptions. They turn out to be valid in many theories with small couplings, but they are also known to be broken in theories with large couplings. Not

surprisingly, perturbation theory does not work very well at all at strong coupling, and the physical states might not correspond to fields in the Lagrangian (or even be particle-like).

The first things we will attempt to compute are the n -point functions, the expectation values of time-ordered products of fields. For this, we use the obvious generalization of the master formula:

$$\langle \Omega | T \{ \mathcal{O}(x, y, z) \} | \Omega \rangle = \frac{\langle 0 | T \{ \mathcal{O}_I(x, y, z) \exp \left[-i \int d^4 w \Delta \mathcal{H}_I(w) \right] \} | 0 \rangle}{\langle 0 | T \{ \exp \left[-i \int d^4 w \Delta \mathcal{H}_I(w) \right] \} | 0 \rangle}, \quad (8.2)$$

where $\mathcal{O}(x, y, z)$ is a shorthand for

$$\mathcal{O}(x, y, z) := \phi(x_1) \dots \phi(x_\ell) \Psi_{a_1}(y_1) \dots \Psi_{a_m}(y_m) \bar{\Psi}^{b_1}(z_1) \dots \bar{\Psi}^{b_n}(z_n). \quad (8.3)$$

Note that the quantities on the left-hand side of Eq. (8.2) refer to the vacuum and the Heisenberg-picture fields of the full theory, while the quantities on the right-hand side involve the vacuum of H_0 and the interaction-picture fields. Furthermore, the endpoints of the time integration within the exponentials should formally be viewed as the limits of $w^0 \rightarrow \pm\infty(1 - i\epsilon)$. All these features are identical to the bosonic case, and the derivation goes through in the same way. The only substantial difference is that the time ordering picks up an extra minus sign whenever two fermion fields are interchanged.

8.1.2 Wick's Theorem

We apply the master formula of Eq. (8.2) to compute n -point functions by expanding the exponentials to a fixed order in the couplings and evaluating the vacuum expectation value of the resulting products of fields. A useful result for doing so is Wick's theorem as generalized to fermions. The result is (for free fields but also applicable to interaction-picture fields)

$$T \{ \mathcal{O}(x, y, z) \} = N \{ \mathcal{O}(x, y, z) + \text{all contractions} \}. \quad (8.4)$$

This is identical to what we had for the scalar theory, although we must still define what we mean by a contraction. The correct definitions are:

$$\overbrace{\phi(x)\phi(x')} = D_F(x - x') \quad (8.5)$$

$$\overbrace{\Psi_a(x)\bar{\Psi}^b(x')} = [S_F(x - x')]_a^b = -\overbrace{\bar{\Psi}^b(x')\Psi_a(x)} \quad (8.6)$$

$$\overbrace{\Psi(x)\Psi(x')} = 0 = \overbrace{\bar{\Psi}(x)\bar{\Psi}(x')} \quad (8.7)$$

$$\overbrace{\phi(x)\bar{\Psi}(x')} = 0 = \overbrace{\phi(x)\bar{\Psi}(x')} \quad (8.8)$$

These contractions reflect the structure of the 2-point functions in the free theory, and should not be too surprising. The only thing to remember is to add a minus sign for every time you anticommute a pair of fermionic fields.

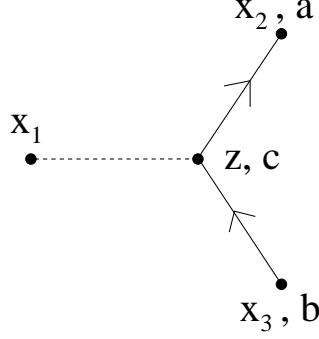


Figure 8.1: Feynman diagram for the $\phi\Psi\bar{\Psi}$ 3-point function at leading order.

8.1.3 Feynman Rules

We are now ready to formulate Feynman rules for n -point functions in perturbation theory. It is obvious from our generalization of Wick's theorem that the 2-point functions at leading order are identical to those of the free theory. Beyond this, we must deduce a vertex factor, and also figure out how to handle multiple fermion lines. As before, we will begin by assigning a Feynman diagram to each distinct contraction contributing to the n -point function. Once we've seen how this works, we will go in the other direction and spell out how to compute the contractions using Feynman diagrams alone.

Let us compute the 3-point function $\phi(x_1)\Psi_a(x_2)\bar{\Psi}^b(x_3)$ to leading order in the coupling y . We have a single, unique contraction:

$$\langle\phi_1\Psi_{2a}\bar{\Psi}_3^b\rangle = \langle 0|T\{\phi_1\Psi_{2a}\bar{\Psi}_3^b(-iy)\int d^4z\phi_z\bar{\Psi}_z^c\Psi_{zc}\}|0\rangle \quad (8.9)$$

$$= (-iy)\int d^4z D_F(x_1-z)[S_F(x_2-z)]_a^c[S_F(z-x_3)]_c^b. \quad (8.10)$$

To do the fermion contractions, $\bar{\Psi}_3^b$ can be moved first all the way to the right, giving a factor of $(-1)^2$, and then everything can be connected up.

The Feynman diagram that we assign to this result is shown in Fig. 8.1. The dashed line denotes the scalar propagator factor D_F , and the solid lines correspond to the fermion propagators S_F . We have also assigned arrows to the fermion lines to show the direction of “index flow”, since $[S_F]_a^b$ is a matrix connecting the index a to the index b . With our choice of conventions, this matches up conveniently with the flow of fermion number. By following the lines backwards, we automatically pick up the correct Dirac index structure, with a connecting to c and c connecting to b . Note as well that the internal index c gets summed over.

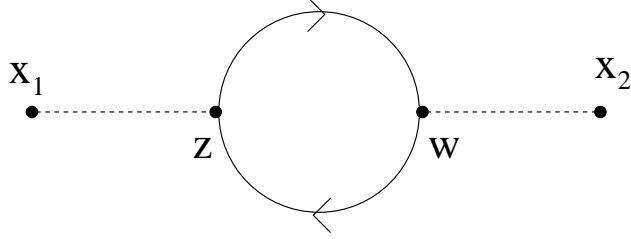


Figure 8.2: Feynman diagram for the connected part of the $\phi\phi$ 2-point function at y^2 order.

With an eye on scattering, let us take the Fourier transform of this result,

$$\begin{aligned} & \int d^4x_1 e^{-ip_1 \cdot x_1} \int d^4x_2 e^{ip_2 \cdot x_2} \int d^4x_3 e^{-ip_3 \cdot x_3} \langle \phi_1 \Psi_{2a} \bar{\Psi}_3^b \rangle \\ &= (-iy) (2\pi)^4 \delta^{(4)}(p_2 - p_1 - p_3) \frac{i}{p_1^2 - M^2} \left[\frac{i(\not{p}_2 + m)}{p_2^2 - m^2} \frac{i(\not{p}_3 + m)}{p_3^2 - m^2} \right]_a^b. \end{aligned} \quad (8.11)$$

You can probably already see how the momentum-space Feynman rules are going to turn out. As before, the Dirac indices contract by following the fermion lines backwards. Note that I have also chosen an opposite sign in the exponential for p_2 relative to the others. This choice corresponds to p_2 flowing out, while p_1 and p_3 both flow in. It is convenient in this case because the momentum flow matches up with the fermion number flow. This doesn't always have to be the case.

As a second example, consider the connected y^2 corrections to the scalar 2-point function. These give

$$\begin{aligned} & \langle \Omega | T \{ \phi(x_1) \phi(x_2) \} | \Omega \rangle \\ &= -(-iy)^2 \int d^4z \int d^4w D_F(z - x_1) [S_F(w - z)]_a^b [S_F(z - w)]_b^a D_F(x_2 - w) \end{aligned} \quad (8.12)$$

$$= -(-iy)^2 \int d^4z \int d^4w D_F(z - x_1) D_F(x_2 - w) \text{tr}[S_F(z - w) S_F(w - z)]. \quad (8.13)$$

The corresponding Feynman diagram is shown in Fig. 8.2. We see that the fermion loop produces a trace over the Dirac indices. This loop also generates a factor of (-1) from the fermion rearrangements needed to contract them all. Both features always occur for closed fermion loops. In momentum space, the result is

$$\begin{aligned} & \int d^4x_1 e^{-ip_1 \cdot x_1} \int d^4x_2 e^{ip_2 \cdot x_2} \langle \phi_1 \phi_2 \rangle \\ &= -(-iy)^2 \left(\frac{i}{p_1^2 - M^2} \right)^2 \int \frac{d^4q}{(2\pi)^4} \text{tr} \left[\frac{i(\not{p}_1 + \not{q} + m)}{(p_1 + q)^2 - m^2} \frac{i(\not{q} + m)}{q^2 - m^2} \right] (2\pi)^4 \delta^{(4)}(p_1 - p_2). \end{aligned} \quad (8.14)$$

Using the gamma matrix tricks discussed in [notes-07](#), it is straightforward to evaluate the trace and reduce it to a function of q , p_1 , and m .

Based on these results, we define the $(\ell + m + n)$ -point function ($\ell\phi$ fields, $m\Psi$ fields, $n\bar{\Psi}$ fields) in momentum space exactly as before:

$$(2\pi)^4 \delta^{(4)} \left(\sum_{i=\ell, m, n} p_i \right) \tilde{G}^{(\ell+m+n)}(\{p\}) = \left(\prod_{i=\ell, m, n} \int d^4 x_i e^{-ip_i \cdot x_i} \right) G^{(\ell+m+n)}(\{x\}), \quad (8.15)$$

where $G^{(\ell+m+n)}(\{x\})$ is the corresponding quantity in position space obtained from the master formula. While we know how to do this, it can be very tedious. Instead, it is more efficient to calculate $\tilde{G}(\{p\})$ directly using a set of (momentum space) Feynman rules:

1. Draw an external line for each momentum p_i with one free end and one fixed end. Also, for each fermion line corresponding to $\Psi_a(y)$, assign the Dirac index a to the external point and draw an arrow on the line directed at the point. Similarly, for each line corresponding to $\bar{\Psi}^b(z)$ include the Dirac index b at the external point and draw an arrow on the line directed away from the point.
2. Draw another M dots corresponding to the vertices, and assign a Dirac index c to each one. Each dot should have one scalar line attached to it, one fermion line directed out of it, and one fermion line directed into it.
3. Assemble all possible Feynman diagrams by connecting up the free ends of lines in pairs in all possible ways subject to two rules:
 - a) Scalar lines can only connect to scalar lines.
 - b) Fermion lines can only be connected if their arrows point in the same direction.
 Also, do not include diagrams that are related by permuting the labels of the vertices. (The $M!$ such permutations cancel the $1/M!$ from expanding the exponential.)
4. Remove all diagrams containing vacuum bubbles and any diagrams with one or more unconnected free ends.
5. Assign a value to each diagram:
 - a) Assign a momentum to every line. Each line connected to an external point gets a four momentum p_i directed inwards (away from the point). Having fixed these, constrain the momenta of all internal lines by imposing four-momentum conservation at every vertex. This may still leave a few internal momenta undertermined. Call them q_j for now.
 - b) Each scalar line with momentum p gets a propagator factor of $i/(p^2 - M^2 + i\epsilon)$. Each fermion line with the momentum flowing parallel to the fermion number direction gets a factor of $i(\not{p} + m)_a^b/(p^2 - m^2 + i\epsilon)$, where a is the Dirac index at the tip of the line and b is Dirac index at the tail (as determined by the direction of the arrow on the line). Also, if the momentum is antiparallel to the fermion number arrow, flip the sign of the momentum in the propagator $p \rightarrow -p$ so that the line gets a factor of $i(-\not{p} + m)_a^b/(p^2 - m^2 + i\epsilon)$.
 - c) Write a factor of $-iy$ for each vertex.
 - d) Integrate over all undetermined momenta $\int d^4 q_j / (2\pi)^4$ and sum over all internal

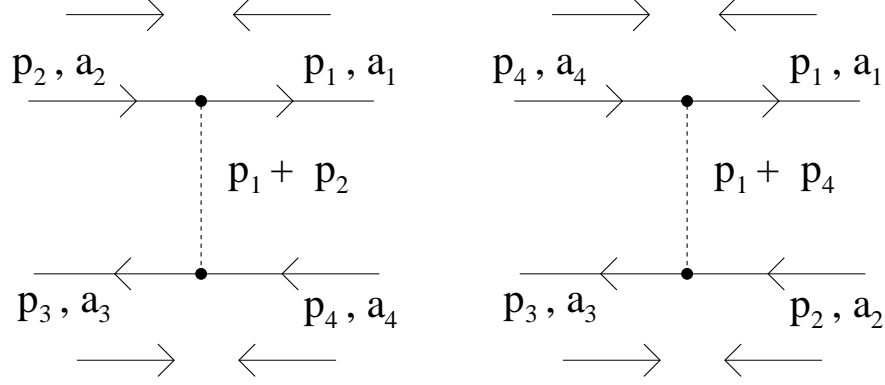


Figure 8.3: Feynman diagram for the connected part of the $\Psi\bar{\Psi}\Psi\bar{\Psi}$ 4-point function.

Dirac indices c .

e) For each set of external momenta p_i connected to each other in some way, multiply by an overall factor of $(2\pi)^4\delta^{(4)}(\sum p_i)$.

f) Multiply the diagram by the symmetry factor and whatever factors of (-1) that were incurred by moving fermions around in the contractions. In general, a closed fermion loop always picks up a net factor of (-1) , and any two diagrams that differ by the exchange of fermion legs have a relative sign between them.

The resulting sum of all the diagrams is the order y^M contribution to $(2\pi)^4\delta^{(4)}(\sum p)\tilde{G}(p)$.

To illustrate these rules in action, let us compute the two diagrams shown in Fig. 8.3, that occur in the y^2 contribution to $\langle\Psi_{a_1}(x_1)\bar{\Psi}_{a_2}(x_2)\Psi_{a_3}(x_3)\bar{\Psi}_{a_4}(x_4)\rangle$. For the first (on the left), we get

$$D_1 = (-iy)^2 \frac{i}{(p_1 + p_2)^2 - M^2} (2\pi)^4\delta^{(4)}\left(\sum_{i=1}^4 p_i\right) \quad (8.16)$$

$$\times \frac{i(-\not{p}_1 + m)_{a_1}^c}{p_1^2 - m^2} \frac{i(\not{p}_2 + m)_c^{a_2}}{p_2^2 - m^2} \frac{i(-\not{p}_3 + m)_{a_3}^d}{p_3^2 - m^2} \frac{i(\not{p}_4 + m)_d^{a_4}}{p_4^2 - m^2},$$

while for the second diagram (on the right) we find

$$D_2 = (-1)(-iy)^2 \frac{i}{(p_1 + p_4)^2 - M^2} (2\pi)^4\delta^{(4)}\left(\sum_{i=1}^4 p_i\right) \quad (8.17)$$

$$\times \frac{i(-\not{p}_1 + m)_{a_1}^c}{p_1^2 - m^2} \frac{i(\not{p}_4 + m)_c^{a_4}}{p_4^2 - m^2} \frac{i(-\not{p}_3 + m)_{a_3}^d}{p_3^2 - m^2} \frac{i(\not{p}_2 + m)_d^{a_2}}{p_2^2 - m^2}.$$

Note the relative minus sign in D_2 relative to D_1 . Again, we see that the general trick to figuring out the Dirac contraction structure is to simply work backwards along the fermion lines.

8.2 Asymptotic States and Scattering

Our next challenge is to convert our Feynman rules for n -point functions into amplitudes for particle scattering. The procedure for this goes through just like for scalars but with a few twists. We will not go into the details, but we will present the main results.

8.2.1 Spectral Decomposition

The spectral decomposition for an interacting Dirac fermion is

$$\int d^4x e^{-ip \cdot x} \langle \Psi_a(x) \bar{\Psi}^b(0) \rangle = \frac{iZ(\not{p} + m)_a^b}{p^2 - m^2 + i\epsilon} + \int_{>m^2}^{\infty} \frac{ds}{2\pi} \frac{W_a^b(s)}{p^2 - s + i\epsilon}, \quad (8.18)$$

where the first term comes from our assumption about the existence of a one-particle state and W_a^b is some unspecified function of s in the Dirac space where we dump everything else. The derivation of this result goes through just like in the scalar case. Note that the location of the pole $p^2 = m^2$ may differ from the mass parameter in the original Lagrangian. The key thing to take away from this result is that an isolated one-particle state corresponds to a pole in the complex p^0 plane.

8.2.2 LSZ for Fermions

We will apply this result to produce a LSZ formula for external fermion states. To motivate the formula, it will be instructive to compute a few expectation values in the free theory. After working in the free theory for a bit, we will come back to the full interacting theory.

Using our definition of states in the free theory, we find the inner product of one-particle states to be

$${}_a \langle k, s | p, r \rangle_a = (2\pi)^3 2p^0 \delta^{rs} \delta^3(\vec{k} - \vec{p}), \quad (8.19)$$

where $|k, s\rangle_a = a^\dagger(k, s)|0\rangle$, and similarly for b -type particles. We also find the free-theory matrix elements

$$\langle 0 | \Psi_c(x) | p, s \rangle_a = u_c(p, s) e^{-ip \cdot x} \quad (8.20)$$

$$\langle 0 | \bar{\Psi}^d(x) | p, s \rangle_b = \bar{v}^d(p, s) e^{-ip \cdot x} \quad (8.21)$$

$${}_b \langle p, s | \Psi_c(x) | 0 \rangle = v_c(p, s) e^{ip \cdot x} \quad (8.22)$$

$${}_a \langle p, s | \bar{\Psi}^d(x) | 0 \rangle = \bar{u}^d(p, s) e^{ip \cdot x}. \quad (8.23)$$

The subscripts on the one-particle states refer to whether they are a -type (particles) or b -type (antiparticles). All the other possible combinations vanish. A handy way to think about these matrix elements is that we have contracted the field with the one-particle state.

An immediate implication of these matrix elements is that (in the free theory)

$${}_a \langle p_1, r | \left[\int d^4z \bar{\Psi}^c(z) \Psi_c(z) \right] | p_2, s \rangle_a = (2\pi)^4 \delta^{(4)}(p_1 - p_2) \bar{u}^c(p_1, r) u_c(p_2, s) + \dots, \quad (8.24)$$

where the omitted terms have $\Psi_c(z)$ and $\bar{\Psi}^c(z)$ contracted with each other. Let's compare this to the Fourier transform of the connected part of $\langle \Psi_{1a} \bar{\Psi}_2^b \int d^4 z \bar{\Psi}_z^c \Psi_{zc} \rangle$:

$$FT \langle \Psi_{1a} \bar{\Psi}_2^b \int d^4 z \bar{\Psi}_z^c \Psi_{zc} \rangle_c = \int d^4 x_1 e^{ip_1 \cdot x_1} \int d^4 x_2 e^{-ip_2 \cdot x_2} \langle \Psi_{1a} \bar{\Psi}_2^b \int d^4 z \bar{\Psi}_z^c \Psi_{zc} \rangle \quad (8.25)$$

$$= (2\pi)^4 \delta^{(4)}(p_1 - p_2) \frac{i(\not{p}_1 + m)_a^c}{p_1^2 - m^2 + i\epsilon} \frac{i(\not{p}_2 + m)_c^b}{p_2^2 - m^2 + i\epsilon}. \quad (8.26)$$

These are clearly different. However, suppose we multiply Eq. (8.26) by $(p_1^2 - m^2 + i\epsilon) \bar{u}^a(p_1, r)$ and $(p_2^2 - m^2 + i\epsilon) u_b(p_2, s)$, and sum over the a and b indices. Using

$$(\not{p} + m)_a^c = \sum_{r'} u_a(p, r') \bar{u}^c(p, r'), \quad \text{and} \quad \bar{u}^a(p, r) u_a(p, r') = 2m \delta^{rr'}, \quad (8.27)$$

this gives

$$\begin{aligned} & \frac{(-i)^2}{(2m)^2} (p_1^2 - m^2 + i\epsilon) \bar{u}^a(p_1, r) (p_2^2 - m^2 + i\epsilon) u_b(p_2, s) FT \langle \Psi_{1a} \bar{\Psi}_2^b \int d^4 z \bar{\Psi}_z^c \Psi_{zc} \rangle \\ &= (2\pi)^4 \delta^{(4)}(p_1 - p_2) \bar{u}^c(p_1, r) u_c(p_2, s), \end{aligned} \quad (8.28)$$

which is precisely the result of Eq. (8.24). In the free theory, at least, we now see how to connect the time-ordered products of field operators to particle states. If we had wanted antiparticle states instead of particle states, we would have instead multiplied by v^a and \bar{v}^b as the projectors. For multiple initial- or final-state particles, we would have used more field operators and more projectors.

We can now give the LSZ reduction formula for fermions within the interacting theory. Instead of writing a big formula, it will be easier to state it as a series of operations. To compute the connected part of the scattering amplitude with m_a fermions and m_b antifermions in the initial state (with initial momenta $\{k_i\}$ and spins $\{s_i\}$) and n_a fermions and n_b antifermions in the final state (with final momenta $\{p_j\}$ and spins $\{r_j\}$):

1. Compute the connected part of the time-ordered vacuum expectation value of $(m_a + n_b)$ $\bar{\Psi}$ fields and $(m_b + n_a)$ Ψ fields using the master formula (or any other method you can think of).
2. Take the Fourier transform with respect to all the spatial coordinates. Use $\int \frac{d^4 k_i}{(2\pi)^4} e^{-ik_i \cdot x_i}$ for the incoming momenta and $\int \frac{d^4 p_j}{(2\pi)^4} e^{+ip_j \cdot x_j}$ for the outgoing momenta.
3. For each external state, multiply by the appropriate projector:
 - $-i(k^2 - m^2) \left(\frac{1}{2m\sqrt{Z}} \right) u_{a_i}(k_i, s_i)$ for each incoming fermion.
 - $-i(p^2 - m^2) \left(\frac{1}{2m\sqrt{Z}} \right) \bar{u}^{b_j}(p_j, r_j)$ for each outgoing fermion.
 - $-i(k^2 - m^2) \left(\frac{1}{2m\sqrt{Z}} \right) \bar{v}^{b_i}(k, s_i)$ for each incoming antifermion.
 - $-i(p^2 - m^2) \left(\frac{1}{2m\sqrt{Z}} \right) v_{a_j}(p_j, r_j)$ for each outgoing antifermion.

Here, Z refers to the factor appearing in the one-particle portion of the fermion spectral formula, Eq. (8.18). Also, the Dirac indices on the spinors should match those of the corresponding fields and be summed over.

4. Take the limits $k_i^2 \rightarrow m^2$ and $p_i^2 \rightarrow m^2$.

The final result is the matrix element $\langle \{p_j, r_j\} | \{k_i, s_i\} \rangle_c$ in the interacting theory. Aside from the factors of \sqrt{Z} (which are equal to unity at leading order in the perturbative expansion), these steps match what we did in the free theory. The new non-trivial physics content is that contributions from well-separated interacting particles in the initial and final states can still be identified with the poles in the n -point functions of the interacting theory.

To illustrate this procedure, let us apply it to the matrix element $\langle \phi(x_1) \Psi_a(x_2) \bar{\Psi}^b(x_3) \rangle$ we computed previously. Looking at Eq. (8.11), the signs we used in the Fourier transform correspond to p_1 and p_3 incoming and p_2 outgoing. Given our LSZ formula and the field content of the matrix element, this is the correct combination for an incoming scalar (p_1) and fermion (p_3), and an outgoing fermion (p_2). Applying the projectors, we get (to leading order)

$$\langle p_2 | p_1 p_3 \rangle_c = (-iy) \frac{1}{Z_\Psi \sqrt{Z_\phi}} (2\pi)^4 \delta^{(4)}(p_2 - p_1 - p_3) \bar{u}^c(p_2, r_2) u_c(p_3, s_3) . \quad (8.29)$$

Note that $Z_\Psi = Z_\phi = 1$ at this order, but we have written them here for posterity. By changing signs in the Fourier transforms and using different projectors, we could have also extracted from this matrix element the amplitudes for incoming and outgoing antifermions, or a fermion-antifermion pair in the initial or final state.

8.2.3 Feynman Rules for Scattering

While the LSZ formula is important conceptually and gives an algorithm to compute scattering matrix elements, it is much easier to compute them directly with a set of Feynman rules. These Feynman rules are very similar to those we formulated in momentum space for n -point functions, but with a different prescription for external legs.

The Feynman rules for our scalar-fermion theory are:

1. Draw all possible connected diagrams at whatever order in the coupling y you are interested in following the same rules as before. The only difference is in external fermion legs. For every incoming or outgoing fermion line, the direction of the line should follow the direction of the momentum flow. However, for every incoming or outgoing antifermion line, the direction of the line is *opposite* the direction of momentum flow. See Fig. 8.4 for an illustration. We do not distinguish between fermions and antifermions in internal lines, but the directions of the fermion lines should always match up.
2. Assign a value to every diagram:
 - a) For all the internal pieces, everything goes through as before.

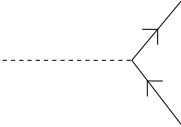
Incoming Fermion	$s \xrightarrow{\mathbf{p}} \bullet$	$u(\mathbf{p},s)$
Incoming Anti-Ferm	$s \xleftarrow{\mathbf{p}} \bullet$	$\bar{v}(\mathbf{p},s)$
Outgoing Fermion	$\bullet \xrightarrow{\mathbf{p}} s$	$\bar{u}(\mathbf{p},s)$
Outgoing Anti-Ferm	$\bullet \xleftarrow{\mathbf{p}} s$	$v(\mathbf{p},s)$
Incoming Scalar	$\cdots \xrightarrow{\mathbf{p}} \bullet$	1
Outgoing Scalar	$\bullet \xleftarrow{\mathbf{p}} \cdots$	1
Internal Fermion	$\bullet \xrightarrow{\mathbf{p}} \bullet$	$i(\mathbf{p} + m)/(\mathbf{p}^2 - m^2)$
Internal Scalar	$\bullet \cdots \bullet$	$i/(\mathbf{p}^2 - m^2)$
Vertex		$-iy$

Figure 8.4: Feynman rules for the scalar-fermion theory.

- b) For the external legs, follow the prescription given in Fig. 8.4.
- c) To get the Dirac index structure right, go backwards along each fermion line and contract the Dirac indices along the way. All these indices should be contracted in the final result.
- d) Multiply by $1/\sqrt{Z_\Psi}$ for every external fermion and $1/\sqrt{Z_\phi}$ for every external scalar.
- e) Figure out the relative sign of the diagram by looking at the contraction structure. This turns out to be (-1) for every internal fermion loop and an additional relative sign for any two diagrams that differ by the exchange of fermion legs.

The result of all this is the scattering amplitude $-i\mathcal{M}$ with the fixed incoming momenta and spins $\{(k_i, s_i)\}$ and the outgoing momenta and spins $\{(p_j, r_j)\}$.

Chapter 9

Quantum Electrodynamics

Quantum Electrodynamics (QED), finally.

9.1 QED Lite

To get warmed up, let us first do some real calculations in the scalar fermion theory we studied previously,

$$\mathcal{L} = \bar{\Psi} i \gamma^\mu \partial_\mu \Psi - m \bar{\Psi} \Psi + \frac{1}{2} (\partial\phi)^2 - \frac{1}{2} M^2 \phi^2 - y \phi \bar{\Psi} \Psi . \quad (9.1)$$

We will see that this theory is analagous to QED, with Ψ representing the electron and the scalar playing a similar role to the photon.

e.g. 1. Decay of the Scalar

The easiest physical process to calculate in this theory is the decay of the scalar to a fermion-antifermion pair, which is allowed by the interaction of Eq. (9.1) provided $M > 2m$. The relevant Feynman diagram (with time going from left to right) is shown in Fig. 9.1. Applying the Feynman rules, we get the amplitude for $p_0 \rightarrow (p_1, s_1) + (p_2, s_2)$

$$-i\mathcal{M} = -iy \bar{u}^c(p_2, s_2) v_c(p_1, s_1) . \quad (9.2)$$

The quantity we are usually interested in is the total decay width, where we sum over the partial decay widths to all possible final momenta and spins. Summing over spins is also the appropriate thing to do if our detector is unable to measure the spin of the outgoing fermions. Using our formula for the decay rate, we find

$$\Gamma = \frac{1}{2M} \int \frac{d^3 p_1}{2E_1 (2\pi)^3} \int \frac{d^3 p_2}{2E_2 (2\pi)^3} (2\pi)^4 \delta^{(4)}(p_0 - p_1 - p_2) \text{“} |\mathcal{M}|^2 \text{”} , \quad (9.3)$$

where

$$\text{“} |\mathcal{M}|^2 \text{”} = \sum_{s_1=1}^2 \sum_{s_2=1}^2 |\mathcal{M}|^2 . \quad (9.4)$$

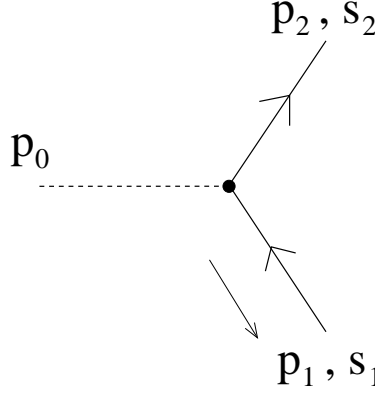


Figure 9.1: Feynman diagram for the decay of the scalar.

Having to sum over spins might seem tedious, but it will actually make our lives easier.¹ Our first task, however, is to take the complex conjugate of \mathcal{M} . We have

$$\mathcal{M}^*/y = [\bar{u}_2 v_1]^* = [u_2^\dagger \gamma^0 v_1]^\dagger = v_1^\dagger (\gamma^0)^\dagger u_2 = \bar{v}_1 u_2 . \quad (9.5)$$

where we have used $(\gamma^0)^\dagger = \gamma^0$ and we have written this in matrix notation. Putting this into the spin sum and writing the Dirac indices explicitly,

$$\frac{1}{y^2} \sum_{s_1, s_2} |\mathcal{M}|^2 = \sum_{s_1, s_2} \bar{v}_1^a u_{2a} \bar{u}_2^b v_{1b} \quad (9.6)$$

$$= \sum_{s_2} u_{2a} \bar{u}_2^b \sum_{s_1} v_{1b} \bar{v}_1^a \quad (9.7)$$

$$= (\not{p}_2 + m)_a^b (\not{p}_1 - m)_b^a \quad (9.8)$$

$$= \text{tr} [(\not{p}_2 + m)(\not{p}_1 - m)] \quad (9.9)$$

$$= 4(p_1 \cdot p_2) - 4m^2 . \quad (9.10)$$

In going to the second line we just rearranged the spinors, and in going to the third line we used have used $\sum_s u(p, s) \bar{u}(p, s) = (\not{p} + m)$ and $\sum_s v(p, s) \bar{v}(p, s) = (\not{p} - m)$. The Dirac index structure of the third line is precisely a trace, which is how we got the fourth line. And for the fifth line, we have used our Dirac trace tricks. This result is representative of what happens in general when we sum over spins – the Dirac structure reduces to one or more traces in the Dirac space. With this “summed and squared” amplitude “ $|\mathcal{M}|^2$ ”, it is just a matter of doing some integrals to get the total decay rate Γ .

e.g. 2. Fermion-Fermion Scattering

As a second example, consider the elastic scattering of a pair of fermions mediated by the scalar. At leading non-trivial order, there are two distinct contributions to this process,

¹ To find the decay rate to specific final spin states, just plug in the explicit forms of the \bar{u} and v spinors.

shown in Fig. 9.2. Each diagram gives its own contribution to the amplitude, and we must sum them all up. The result is

$$\begin{aligned}
-i\mathcal{M} &= -i\mathcal{M}_t - i\mathcal{M}_u & (9.11) \\
&= (-iy)^2 \left[\frac{i}{(p_1 - p_3)^2 - M^2} \bar{u}_3^a u_{a1} \bar{u}_4^b u_{b2} - \frac{i}{(p_1 - p_4)^2 - M^2} \bar{u}_4^a u_{a1} \bar{u}_3^b u_{b2} \right],
\end{aligned}$$

where u_{a1} stands for $u_a(p_1, s_1)$ and so on. The relative minus sign comes from the contraction structure, much like in the related example discussed in **notes-08**. Note that the two diagrams differ by the exchange of the fermions 3 and 4.

When we square the amplitude to get $|\mathcal{M}|^2$ there will be interference between the two contributions leading to four distinct terms. This can get cumbersome, but we can use the same methods as before to evaluate them. Putting this result into our formula for scattering cross sections, we can find the total (or differential) scattering cross section for a given values of the initial spins s_1 and s_2 the final spins s_3 and s_4 . However, most scattering experiments are unpolarized, in that they do not have fixed values for the spins of the initial states. If this is the case, we should *average* the cross section over the spins of the initial states. Furthermore, if the experiment does not measure the spins of the final states, we should also sum the cross section over final spins. The expression for the *total unpolarized cross section* in this case is therefore given by the same expression we had previously for the total cross section, but with

$$|\mathcal{M}|^2 \rightarrow \text{“}|\mathcal{M}|^2\text{”} = \frac{1}{4} \sum_{s_1, s_2} \sum_{s_3, s_4} |\mathcal{M}|^2, \quad (9.12)$$

where the factor of $1/4$ comes from averaging over the $4 = 2 \times 2$ possible initial spins.

To illustrate this, let's compute a few of the terms that arise. We have

$$|\mathcal{M}|^2 = |\mathcal{M}_t|^2 + \mathcal{M}_t \mathcal{M}_u^* + \mathcal{M}_t^* \mathcal{M}_u + |\mathcal{M}_u|^2. \quad (9.13)$$

The first term gives

$$\frac{(t - M^2)^2}{y^4} \times \sum_{s, s'} |\mathcal{M}_t|^2 = \sum_{\{s\}} (\bar{u}_3^a u_{1a} \bar{u}_1^b u_{3b}) (\bar{u}_4^c u_{2c} \bar{u}_2^d u_{4d}) \quad (9.14)$$

$$= \text{tr} [(\not{p}_3 + m)(\not{p}_1 + m)] \text{tr} [(\not{p}_4 + m)(\not{p}_2 + m)] \quad (9.15)$$

$$= 16 [(p_1 \cdot p_3) + m^2] [(p_2 \cdot p_4) + m^2]. \quad (9.16)$$

The second term is a bit trickier. Here, it really helps to write out the Dirac indices:

$$\frac{(t - M^2)(u - M^2)}{y^4} \times \sum_{\{s\}} \mathcal{M}_t \mathcal{M}_u^* = - \sum_{\{s\}} (\bar{u}_3^a u_{1a} \bar{u}_4^b u_{2b}) (\bar{u}_1^c u_{4c} \bar{u}_2^d u_{3d}) \quad (9.17)$$

$$= -(\not{p}_1 + m)_a^c (\not{p}_4 + m)_c^b (\not{p}_2 + m)_b^d (\not{p}_3 + m)_d^a \quad (9.18)$$

$$= -\text{tr} [(\not{p}_1 + m)(\not{p}_4 + m)(\not{p}_2 + m)(\not{p}_3 + m)] \quad (9.19)$$

It is straightforward to compute this trace using our trace tricks and I will leave it to you. You should now also be able to compute the other two terms that contribute to the summed and squared amplitude.

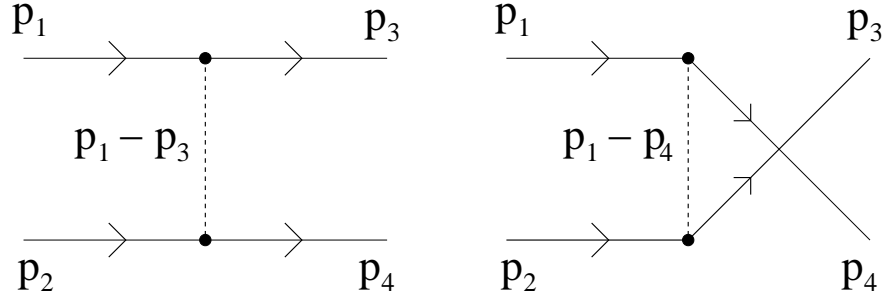


Figure 9.2: Feynman diagram for the elastic scattering of fermions.

9.2 QED

After this warmup, let us now present the Lagrangian for QED. It is

$$\mathcal{L} = \bar{\Psi} i \gamma^\mu D_\mu \Psi - m \bar{\Psi} \Psi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} , \quad (9.20)$$

where

$$D_\mu = \partial_\mu + ieQA_\mu , \quad (9.21)$$

and

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu . \quad (9.22)$$

That's it! We identify the fermion Ψ with the electron and the vector field A_μ with the photon.² Indeed, the vector term coincides with the Lagrangian we had for classical electromagnetism in the absence of sources.

Our convention is to set $Q = -1$ and to take the coupling e to be the magnitude of the electron charge. In our natural units it is dimensionless. The experimentally measured value is

$$\alpha := \frac{e^2}{4\pi} \simeq 1/(137.036) . \quad (9.23)$$

We will see that our perturbative expansion in the coupling corresponds to an expansion in α . Since $\alpha \ll 1$, this expansion should converge very quickly.

9.2.1 Gauge Invariance

The QED Lagrangian has a global $U(1)$ symmetry under the transformations

$$\begin{cases} \Psi & \rightarrow e^{-iQ\alpha}\Psi \\ A_\mu & \rightarrow A_\mu \end{cases} , \quad (9.24)$$

²We will add other fermions later on.

for any constant α . The corresponding Noether current is

$$j^\mu = -\bar{\Psi}\gamma^\mu\Psi . \quad (9.25)$$

On the other hand, suppose we're feeling adventurous and decide to elevate the transformation parameter to a function on spacetime: $\alpha = \alpha(x)$. Doing so, we find that the transformation above is no longer a symmetry of the Lagrangian. In particular,

$$\bar{\Psi}i\gamma^\mu\partial_\mu\Psi \rightarrow \bar{\Psi}i\gamma^\mu\partial_\mu\Psi + \bar{\Psi}i\gamma^\mu(-iQ\partial_\mu\alpha)\Psi . \quad (9.26)$$

Evidently the transformation of Eq. (9.24) is *not* a symmetry of the theory for non-constant parameters $\alpha(x)$.

We can restore the invariance of the fermion terms if we also have the vector field transform according to:

$$\begin{cases} \Psi & \rightarrow e^{-iQ\alpha}\Psi \\ A_\mu & \rightarrow A_\mu + \frac{1}{e}\partial_\mu\alpha . \end{cases} \quad (9.27)$$

Together, this implies that

$$(\partial_\mu + ieQA_\mu)\Psi \equiv D_\mu\Psi \rightarrow e^{-iQ\alpha}D_\mu\Psi , \quad (9.28)$$

and therefore $\bar{\Psi}i\gamma^\mu D_\mu\Psi$ is invariant under the transformation for arbitrary $\alpha(x)$ provided the vector field also transforms as indicated. The differential operator D_μ is sometimes called a *covariant derivative*. Even better, if we look at the effect of this shift on the photon kinetic term, we find that it remains unchanged as well:

$$F_{\mu\nu} = (\partial_\mu A_\nu - \partial_\nu A_\mu) \rightarrow (\partial_\mu A_\nu - \partial_\nu A_\mu) - \frac{1}{e}(\partial_\mu\partial_\nu - \partial_\nu\partial_\mu)\alpha = F_{\mu\nu} + 0 . \quad (9.29)$$

Thus, QED is invariant under the transformations of Eq. (9.27) for any reasonable arbitrary function $\alpha(x)$.

At first glance this invariance might just seem like a clever trick, but the river beneath these still waters runs deep. Thinking back to regular electromagnetism (of which QED is just the quantized version), one often deals with scalar and vector potentials. These potentials are not unique and are therefore not observable (for the most part), and the true “physical” quantities are the electric and magnetic fields. The invariance of the Lagrangian reflects this fact. These transformations are called *gauge transformations*, and the photon derived from A_μ is sometimes said to be a gauge boson.

Keeping in mind the story from electromagnetism, the interpretation of the quantum fields in QED is that ***only those quantities that are invariant under the transformations of Eq. (9.27) are physically observable***. In particular, the vector field A_μ that represents the photon is not itself an observable quantity, but the gauge-invariant field strength $F_{\mu\nu}$ is. Put another way, the field variables we are using are redundant, and the transformations of Eq.(9.27) represent an *equivalence relation*: any two set of fields (Ψ, A_μ) related by such a transformation represent the same physical configuration. Sometimes

the invariance under Eq. (9.27) is called a local or gauge symmetry, but it is not really a symmetry at all. A true symmetry implies that different physical configurations have the same properties. Gauge invariance is instead a statement about which configurations are physically observable.

Gauge invariance is also sensible if we consider the independent polarization states of the photon, of which there are two. The vector field A_μ represents the photon, but it clearly has four independent components. Of these, the timelike polarization component is already non-dynamical on account of the form of the vector kinetic term. Invariance under gauge transformations effectively removes the additional longitudinal polarization leaving behind only the two physical transverse polarization states. Note as well that if the photon had a mass term, $\mathcal{L} \supset m^2 A_\mu A^\mu / 2$, the theory would no longer be gauge invariant. Instead, the longitudinal polarization mode would enter as physical degree of freedom. Equivalently, gauge invariance forces the photon to be massless.

In the discussion above we started with the QED Lagrangian and showed that it was gauge-invariant. However, the modern view is to take gauge invariance as the fundamental principle. Indeed, the only way we know of to write a consistent, renormalizable theory of interacting vector fields is to have an underlying gauge symmetry. For QED, we could have started with a local $U(1)$ gauge invariance for a charged fermion field and built up the rest of the Lagrangian based on this requirement. In this context, the vector field is needed to allow us to define a sensible derivative operator on the fermion field, which involves taking a difference of two fields at different spacetime points with apparently different transformation properties, and corresponds to something called a *connection*.³ Gauge invariance completely fixes the photon-fermion interactions up to the overall charge Q , illustrating why it is so powerful.

9.2.2 Feynman Rules

It turns out that quantizing the vector field A_μ using the tools we've developed is a bit tricky on account of the gauge redundancy. We will come back to this a bit later. For now, we will simply present the result of the quantization procedure in the form of Feynman rules for the photon and the electron. This is all that is needed to compute scattering processes.

The photon field A_μ has a mode expansion similar to that of the fermion in that every momentum mode has a polarization associated to it. In this case, we need two (possibly complex) polarization basis vectors $\epsilon^\mu(p, \lambda)$, $\lambda = 1, 2$. These states are orthogonal,

$$\epsilon^\mu(p, \lambda) \epsilon_\mu^*(p, \lambda') = -\delta^{\lambda\lambda'} , \quad (9.30)$$

and they satisfy a truncated completeness relation,

$$\sum_\lambda \epsilon^\mu(p, \lambda) \epsilon^{\nu*}(p, \lambda) = -\eta^{\mu\nu} + (p^\mu \text{ stuff}) . \quad (9.31)$$

³ See Ch. 15 of Peskin & Schroeder for a nice explanation of these slightly cryptic comments.

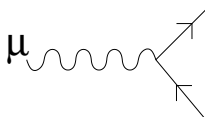
Incoming Fermion	$s \xrightarrow{\mathbf{p}} \bullet$	$u(\mathbf{p},s)$
Incoming Anti-Ferm	$s \xleftarrow{\mathbf{p}} \bullet$	$\bar{v}(\mathbf{p},s)$
Outgoing Fermion	$\bullet \xrightarrow{\mathbf{p}} s$	$\bar{u}(\mathbf{p},s)$
Outgoing Anti-Ferm	$\bullet \xleftarrow{\mathbf{p}} s$	$v(\mathbf{p},s)$
Incoming Photon	$\mu, \lambda \begin{array}{c} \mathbf{p} \\ \text{wavy line} \end{array} \bullet$	$\epsilon_\mu(\mathbf{p}, \lambda)$
Outgoing Photon	$\bullet \begin{array}{c} \mathbf{p} \\ \text{wavy line} \end{array} \mu, \lambda$	$\epsilon_\mu^*(\mathbf{p}, \lambda)$
Internal Fermion	$\bullet \xrightarrow{\mathbf{p}} \bullet$	$i(\mathbf{p} + m)/(\mathbf{p}^2 - m^2)$
Internal Photon	$\mu \begin{array}{c} \mathbf{p} \\ \text{wavy line} \end{array} \nu$	$-i\eta_{\mu\nu}/\mathbf{p}^2$
Vertex		$-ieQ\gamma^\mu$

Figure 9.3: Feynman rules for QED.

We will see shortly that the p^μ stuff can be ignored. In terms of these polarization basis vectors, the mode expansion is

$$A^\mu(x) = \int \widetilde{dk} \sum_\lambda [a(k, \lambda)\epsilon^\mu(k, \lambda)e^{-ik \cdot x} + a^\dagger(k, \lambda)\epsilon^{\mu*}(k, \lambda)e^{ik \cdot x}] , \quad (9.32)$$

where $k^0 = \sqrt{\vec{k}^2}$.

Much like for external fermion states, an outgoing external photon will contribute a factor of $\epsilon_\mu^*(p, \lambda)$ to the amplitude, and an incoming photon will give $\epsilon_\mu(p, \lambda)$. Also like for fermions, the photon propagator takes the form

$$\frac{\sum_\lambda \epsilon^\mu(p, \lambda)\epsilon^{\nu*}(p, \lambda)}{p^2 + i\epsilon} = \frac{i}{p^2 + i\epsilon} (-\eta^{\mu\nu} + p^\mu \text{ stuff}) . \quad (9.33)$$

Again, we will see that the p^μ stuff can be ignored.

Having made these introductions, we can now state the Feynman rules for QED. They are identical to those we had for QED Lite, but with photons instead of scalars. This leads to three essential differences that we summarize in Fig. 9.3. The first is that the photon-fermion-antifermion vertex is

$$\text{Vertex} = -iQe\gamma^\mu . \quad (9.34)$$

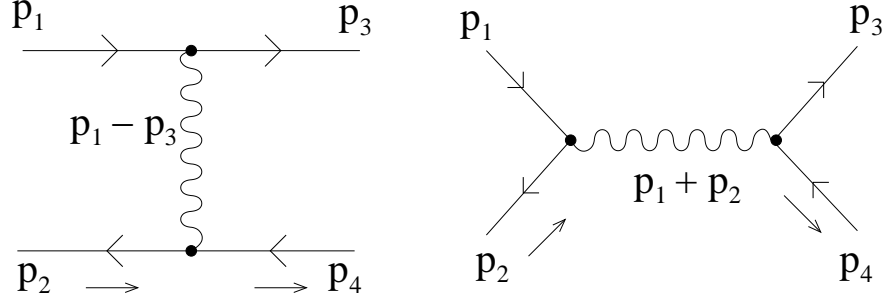


Figure 9.4: Diagrams for Bhabha scattering at leading order.

The second is that we use a photon propagator (dropping the p^μ stuff) instead of the scalar propagator. And third, external photon lines contribute polarization vectors $\epsilon^{(*)}(p, \lambda)$ to the value of the diagram. At the end of the day, all vector indices should be contracted.

9.3 Computing Stuff

Knowing how to compute things in QED is the first step to doing research in particle physics. These skills also transfer over to many other areas of physics. We present here a few specific examples of standard QED processes.

e.g. 3. Electron-Positron (Bhabha) Scattering

The antiparticle of the electron, the antielectron, is sometimes also called the positron. At leading order, there are two diagrams for this process shown in Fig. 9.4. The amplitude is therefore

$$\begin{aligned}
 -i\mathcal{M} &= -i(\mathcal{M}_t + \mathcal{M}_s) \\
 &= (-ieQ)^2 \left[(\bar{u}_3 \gamma^\mu u_1) (\bar{v}_2 \gamma^\nu v_4) \frac{-i\eta_{\mu\nu}}{(p_1 - p_3)^2 + i\epsilon} - (\bar{v}_2 \gamma^\mu u_1) (\bar{u}_3 \gamma^\nu v_4) \frac{-i\eta_{\mu\nu}}{(p_1 + p_2)^2 + i\epsilon} \right].
 \end{aligned} \tag{9.35}$$

The relative sign between the two pieces requires a slightly careful examination of the contraction structure. It is straightforward to square this and sum over spins.

e.g. 4. Compton Scattering

Compton scattering corresponds to electron-photon scattering, $e\gamma \rightarrow e\gamma$. There are two diagrams at leading order, shown in Fig. 9.5. They are equal to

$$\begin{aligned}
 -i\mathcal{M} &= -i(\mathcal{M}_s + \mathcal{M}_u) \\
 &= (-ieQ)^2 \left[\bar{u}_3 \gamma^\mu \frac{i(\not{q} + m)}{q^2 - m^2} \gamma^\nu u_1 + \bar{u}_3 \gamma^\nu \frac{i(\not{k} + m)}{k^2 - m^2} \gamma^\mu u_1 \right] \epsilon_\mu^*(p_4) \epsilon_\nu(p_2),
 \end{aligned} \tag{9.36}$$

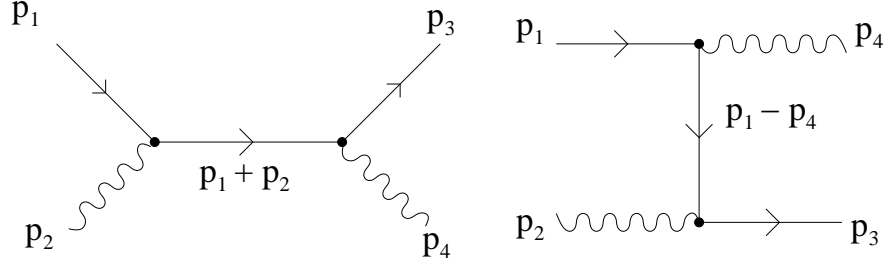


Figure 9.5: Diagrams for Compton scattering at leading order.

where $q = (p_1 + p_2)$ and $k = (p_1 - p_4)$. Note the specific contraction of indices and that we made sure to use the same momentum label for the distinguishable outgoing particles.

To illustrate how to deal with external photons, let us look at the summed and squared amplitude. In many experiments, we only measure the total rate and not the specific polarizations of the outgoing photons (or the fermion spins). Furthermore, the initial beams often have no net polarization. In this case, we should average over initial polarizations and spins and sum over final ones. The summed and squared matrix element that goes into the formula for the total unpolarized cross section is therefore

$$|\mathcal{M}|^{2''} = \frac{1}{2} \frac{1}{2} \sum_{s_1, s_3} \sum_{\lambda_2, \lambda_4} |\mathcal{M}(s_1, s_3, \lambda_2, \lambda_4)|^2, \quad (9.37)$$

where we have divided by the number of initial spins (2) and initial polarizations (2). The s -channel contribution is

$$|\mathcal{M}_s|^{2''} \propto \sum_{s_1, s_3} [\bar{u}_3 \gamma^\mu (\not{q} + m) \gamma^\nu u_1] [\bar{u}_1 \gamma^\beta (\not{q} + m) \gamma^\alpha u_3] \sum_{\lambda_2, \lambda_4} (\epsilon_{\mu 4}^* \epsilon_{\alpha 4}) (\epsilon_{\nu 2} \epsilon_{\beta 2}^*) \quad (9.38)$$

$$= \text{tr}[(\not{p}_3 + m) \gamma^\mu (\not{q} + m) \gamma^\nu (\not{p}_1 + m) \gamma^\beta (\not{q} + m) \gamma^\alpha] (-\eta_{\mu\alpha}) (-\eta_{\nu\beta}). \quad (9.39)$$

This is pretty complicated, but it can be simplified with lots of work or by using clever tricks. A useful one is $\not{p} u(p) = m u(p)$. The t -channel piece goes through similarly.

Chapter 10

Quantizing the Photon

Our remaining task is to justify the photon propagator and external state sums we presented in the Feynman rules for QED. In doing so, we will also gain some intuition about how gauge invariance manifests itself within Feynman diagrams.

10.1 Classical Vector Fields

The classical Lagrangian of electromagnetism (in the absence of sources) is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} , \quad (10.1)$$

where

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu . \quad (10.2)$$

Under Lorentz transformations, the vector field A^μ transforms as

$$A^\mu(x) = \Lambda^\mu{}_\nu A^\nu(\Lambda^{-1}x) , \quad (10.3)$$

corresponding to the 4-vector representation.

Everything looks good so far, but a puzzle arises when we try to quantize the theory. We know that the photon has two independent polarization states, corresponding to the two possible helicities of a massless particle. However, the vector field A^μ has four components, and it would seem that the theory should have four particle states. The solution to this puzzle will turn out to be gauge invariance.

10.1.1 Degrees of Freedom: Massive Vector

Before tackling the photon, let us begin with the slightly easier case of a massive vector field Z^μ . Even though this field would seem to have four degrees of freedom, we can use it to

build a quantum theory of a massive particle with spin $s = 1$, with three degrees of freedom. To see how, note that under the rotation subgroup of Lorentz a general 4-vector decomposes into states with $s = 0$ and $s = 1$. The $s = 0$ piece corresponds to the subset of Z^μ fields that can be written as $Z_\mu = \partial_\mu \phi$ for some scalar ϕ . To project these configurations out to isolate the $s = 1$ part, it turns out to be sufficient to apply the constraint $\partial_\mu Z^\mu = 0$.

To see how this works, let us try using a Lagrangian for Z^μ that is a simple generalization of electromagnetism,

$$\mathcal{L} = -\frac{1}{4}Z_{\mu\nu}Z^{\mu\nu} + \frac{1}{2}m^2Z_\mu Z^\mu \quad (10.4)$$

$$= \frac{1}{2}Z^\mu(\eta_{\mu\nu}\partial^2 - \partial_\mu\partial_\nu)Z^\nu + \frac{1}{2}m^2Z_\mu Z^\mu, \quad (10.5)$$

where $Z_{\mu\nu} = \partial_\mu Z_\nu - \partial_\nu Z_\mu$, and we have integrated by parts in the second line and dropped the total derivative that will vanish when inserted into the action. The only new thing here relative to electromagnetism is the mass term.

The equations of motion implied by this Lagrangian are

$$0 = (\eta_{\mu\nu}\partial^2 - \partial_\mu\partial_\nu)Z^\nu + m^2Z_\mu \quad (10.6)$$

Applying ∂^μ to this result, we find the condition

$$m^2\partial_\mu Z^\mu = 0. \quad (10.7)$$

Thus, the only way to have a non-trivial solution is for the condition $\partial_\mu Z^\mu = 0$ to be satisfied. This is also consistent with the absence of a $s = 0$ mode. Consider what happens in this Lagrangian when $Z_\mu = \partial_\mu \phi$. The kinetic term gives

$$Z_{\mu\nu} \rightarrow \partial_\mu(\partial_\nu\phi) - \partial_\nu(\partial_\mu\phi) = 0. \quad (10.8)$$

In the mass term, we find

$$m^2Z_\mu Z^\mu \rightarrow m^2\partial_\mu\phi\partial^\mu\phi \rightarrow -m^2\phi\partial^2\phi. \quad (10.9)$$

This vanishes as well if we impose the constraint

$$0 = \partial_\mu Z^\mu \rightarrow \partial_\mu(\partial^\mu\phi). \quad (10.10)$$

Thus, the funny kinetic term plus the constraint imply that the action does not depend at all on the $s = 0$ part of Z^μ .

It is also possible to couple the massive vector to other stuff without reintroducing a dependence on the $s = 0$ part. If we were to write an interaction for Z^μ that is linear in the field, it would have to take the form

$$\mathcal{L} \supset Z_\mu j^\mu, \quad (10.11)$$

for some four-vector operator j^μ . Putting $Z_\mu = \partial_\mu \phi$ into Eq. (10.11), we find

$$\mathcal{L} \supset \partial_\mu\phi j^\mu = -\phi(\partial_\mu j^\mu), \quad (10.12)$$

up to a total derivative. This implies that the action will not depend on the $s = 0$ component of Z^μ if and only if the operator j^μ is a conserved current, $\partial_\mu j^\mu = 0$.

With the constraint, the equations of motion for the free massive vector become

$$0 = (\partial^2 + m^2)Z_\mu - \partial_\mu(\partial_\nu Z^\nu) , \quad (10.13)$$

which is just a Klein-Gordon equation for each of the components. A general solution is

$$Z_\mu = \sum_{\lambda=1}^3 \int \widetilde{d^3k} [a(k, \lambda)\epsilon_\mu(k, \lambda)e^{-ik \cdot x} + a^*(k, \lambda)\epsilon_\mu^*(k, \lambda)e^{ik \cdot x}] , \quad (10.14)$$

where the *polarization* 4-vectors ϵ^μ must satisfy

$$k^\mu \epsilon_\mu(k, \lambda) = 0 . \quad (10.15)$$

This condition comes from the $\partial_\mu Z^\mu = 0$ constraint. Three of them are needed to make up a basis of 4-vectors subject to the one constraint. For $k^\mu = (E_k, 0, 0, k)$, a convenient choice for them is

$$\epsilon^\mu(1) = (0, 1, 0, 0), \quad \epsilon^\mu(2) = (0, 0, 1, 0), \quad \epsilon^\mu(3) = \left(\frac{k}{m}, 0, 0, \frac{E_k}{m} \right) . \quad (10.16)$$

The first two are called transverse polarizations, while the third is said to be longitudinal.

10.1.2 Degrees of Freedom: Massless Vector

We turn next to the free photon. The Lagrangian can be rewritten as

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} = \frac{1}{2}A^\mu(\eta_{\mu\nu}\partial^2 - \partial_\mu\partial_\nu)A^\nu , \quad (10.17)$$

where we have again dropped a total derivative. To relate this to a massless particle with two helicity states is more difficult than the massive case, and is not just a matter of taking the mass to zero.

The key new feature of the massless vector Lagrangian of Eq. (10.17) is the invariance under gauge transformations

$$A_\mu \rightarrow A_\mu + \frac{1}{e}\partial_\mu\alpha , \quad (10.18)$$

for any smooth function α . As discussed in [notes-09](#), there is also a new physical interpretation of the theory to go along with this invariance. Specially, field configurations related by a gauge transformation are understood to be physically equivalent. In the language that is typically used in classical electromagnetism, the physical quantities are the electric and magnetic fields, not the scalar and vector potentials, $A^\mu = (\phi, \vec{A})$.

The equations of motion implied by Eq. (10.17) are

$$0 = \partial^2 A^0 - \partial_0(\partial_0 A^0 + \partial_i A^i) = -\vec{\nabla}^2 A^0 - \partial_0(\vec{\nabla} \cdot \vec{A}) \quad (10.19)$$

$$0 = -\partial^2 A^i - \partial_i(\partial_0 A^0 + \partial_j A^j), \quad (10.20)$$

where we have written the time and space components separately. Note that Eq. (10.19) is just a Poisson equation for A^0 . It implies that we can solve for A^0 in terms of the spatial components [18],

$$A^0(\vec{x}) = \frac{1}{4\pi} \int d^3x' \frac{\vec{\nabla} \cdot (\partial \vec{A} / \partial t)}{|\vec{x} - \vec{x}'|}. \quad (10.21)$$

This looks promising for the quantum theory, since it looks like we will be able to remove A^0 as a dynamical variable.

To actually solve the equations of motion for A^μ , it helps enormously to choose a specific gauge. A popular example is the *Coulomb gauge*,

$$\vec{\nabla} \cdot \vec{A} = 0. \quad (10.22)$$

In this case, we get $A^0 = 0$ together with $\partial^2 \vec{A} = \vec{0}$. The general solution is

$$A^i(x) = \sum_{\lambda=1,2} \int \widetilde{dk} \left[a(\vec{k}, \lambda) \epsilon^i(\vec{k}, \lambda) e^{-ik \cdot x} + a^*(\vec{k}, \lambda) \epsilon^{i*}(\vec{k}, \lambda) e^{ik \cdot x} \right], \quad (10.23)$$

with $k^0 = \sqrt{\vec{k}^2}$. Applying the Coulomb gauge condition, we must have $\vec{k} \cdot \vec{\epsilon} = 0$. A convenient choice of basis 3-vectors for $\vec{k} = (0, 0, k)$ are the linear polarizations

$$\epsilon^i(1) = (1, 0, 0), \quad \epsilon^i(2) = (0, 1, 0). \quad (10.24)$$

A second popular choice are the right- and left-handed circular polarizations

$$\epsilon^i(1) = \frac{1}{\sqrt{2}}(1, i, 0), \quad \epsilon^i(2) = \frac{1}{\sqrt{2}}(1, -i, 0). \quad (10.25)$$

In this gauge, we can interpret \vec{A} as a vector wave propagating at the speed of light with two independent polarizations. The downside of the Coulomb gauge is that it obscures the underlying Lorentz invariance of the theory.

A second gauge choice is the Lorentz-invariant *Lorenz gauge*,¹

$$\partial_\mu A^\mu = 0. \quad (10.26)$$

A convenient way to impose this condition on the classical theory is to use a Lagrange multiplier ξ . This means we modify the Lagrangian to

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2\xi} (\partial_\mu A^\mu)^2, \quad (10.27)$$

¹Often called Lorentz gauge, because the corresponding condition is Lorentz invariant. Poor Lorenz.

while treating the A^μ fields as being unconstrained but also promoting ξ to a variable of the system. Without having to worry about the constraint, it is easy to derive the equations of motion for A^μ ,

$$[-\eta_{\mu\nu}\partial^2 + (1 - 1/\xi)\partial_\mu\partial_\nu] A^\nu = 0 . \quad (10.28)$$

However, we also have to include the equation of motion for ξ , which yields $\partial_\mu A^\mu = 0$ and reproduces the constraint.

10.2 Quantizing the Vector

We turn next to quantizing the massless vector. This will turn out to be fairly complicated, with a number of subtleties due to gauge invariance and the related fact that we have more field variables than physical states. We will also work specifically within the Lorenz gauge as in Refs. [18, 19]. See Srednicki [2] for a discussion of photon quantization in the Coulomb gauge.

10.2.1 Hamiltonian Formulation

As usual, the first step is to find the classical Hamiltonian. For now, we will not worry about the gauge choice $\partial_\mu A^\mu = 0$, and just work with Eq. (10.27) as the defining Lagrangian. We will also fix $\xi = 1$. This isn't necessary, but it greatly simplifies the algebra.

The conjugate momenta to the A^μ are

$$\Pi_\mu = \frac{\partial\mathcal{L}}{\partial(\partial_0 A^\mu)} = -F_{0\mu} . \quad (10.29)$$

The Hamiltonian is therefore

$$H = \int d^3x [\Pi_\mu\partial_0 A^\mu - \mathcal{L}] \quad (10.30)$$

$$= \int d^3x \left(\frac{1}{2} [(\pi^i)^2 + (\vec{\nabla} A^i)^2] - \frac{1}{2} [(\pi^0)^2 + (\vec{\nabla} A^0)^2] \right) \quad (10.31)$$

The negative sign in the A^0 piece looks strange, but let's keep on trucking.

10.2.2 Going Quantum

The next step is to elevate the field variables and their conjugate momenta to operators on a Hilbert space. The equal-time commutation relations are

$$[A^\mu(t, \vec{x}), \Pi_\nu(t, \vec{x}')] = i \delta^\mu_\nu \delta^{(3)}(\vec{x} - \vec{x}') , \quad (10.32)$$

$$[A^\mu(t, \vec{x}), A^\nu(t, \vec{x}')] = 0 = [\Pi_\mu(t, \vec{x}), \Pi_\nu(t, \vec{x}')] . \quad (10.33)$$

Note that each of the components of A^μ is treated as an independent variable.

Following the same procedure as for the scalar, we can rewrite the vector field in terms of mode operators,

$$A^\mu(x) = \int \widetilde{dk} \sum_{\lambda=0}^3 [a(k, \lambda) \epsilon^\mu(k, \lambda) e^{-ik \cdot x} + a^\dagger(k, \lambda) \epsilon^{\mu*}(k, \lambda) e^{ik \cdot x}] , \quad (10.34)$$

where $k^0 = |\vec{k}|$, and the four $\epsilon^\mu(k, \lambda)$ form a basis of four vectors. It is convenient (and always possible) to choose these 4-vectors such that ²

$$k \cdot \epsilon(k, 1) = k \cdot \epsilon(k, 2) = 0 \quad (10.35)$$

$$k \cdot \epsilon(k, 0) = -k \cdot \epsilon(k, 3) \quad (10.36)$$

$$\epsilon^\mu(k, \lambda) \epsilon_\mu(k, \lambda') = \eta^{\lambda\lambda'} \quad (10.37)$$

With this choice, the commutation relations of the mode operators implied by the ETCRs are

$$[a_{k,\lambda}, a_{p,\lambda'}^\dagger] = -\eta^{\lambda\lambda'} (2\pi)^3 2k^0 \delta^{(3)}(\vec{k} - \vec{p}) \quad (10.38)$$

$$[a_{k,\lambda}, a_{p,\lambda'}] = 0 = [a_{k,\lambda}^\dagger, a_{p,\lambda'}^\dagger] \quad (10.39)$$

In terms of these modes operators, the Hamiltonian is given by

$$H = \int \widetilde{dk} k^0 \sum_{\lambda\lambda'} (-\eta^{\lambda\lambda'}) a_{k,\lambda}^\dagger a_{k,\lambda'} \quad (10.40)$$

$$= \int \widetilde{dk} k^0 \left[\sum_{\lambda=1}^3 a_{k,\lambda}^\dagger a_{k,\lambda} - a_{k,0}^\dagger a_{k,0} \right] . \quad (10.41)$$

To build the Hilbert space, we assume there exists a vacuum state $|0\rangle$ such that

$$a_{k,\lambda}|0\rangle = 0 . \quad (10.42)$$

All other states in the space can be built by applying powers of $a_{k,\lambda}^\dagger$ to $|0\rangle$. We interpret these states as free photons with four-momentum k^μ and polarization λ .

10.2.3 Problems with this Quantum Theory

As it stands, the quantum theory we've just developed has a number of puzzling features. We collect here a list of the worst of them. In the next subsection, we will present a way to fix them.

~~i) One-Particle States~~

²For example, if $k^\mu = (k, 0, 0, k)$, $\epsilon(0) = (1, 0, 0, 0)$, $\epsilon(1) = (0, 1, 0, 0)$, $\epsilon(2) = (0, 0, 1, 0)$, $\epsilon(3) = (0, 0, 0, 1)$.

This theory has four one-particle states, corresponding to $|k, \lambda\rangle = a_{k,\lambda}^\dagger|0\rangle$ for $\lambda = 0, 1, 2, 3$. However, we know that a real photon has only two independent polarizations. This isn't too surprising because we haven't used gauge invariance yet.

ii) Commutators

The commutator of Eq. (10.38) looks just like what we would expect for independent scalar fields for $\lambda = 1, 2, 3$, but it has the opposite sign for $\lambda = \lambda' = 0$.

iii) Inner Products

Consider the inner product of a pair of one-particle states,

$$\langle k, \lambda|p, \lambda'\rangle = -\eta^{\lambda\lambda'}(2\pi)^3 2k^0 \delta^{(3)}(\vec{k} - \vec{p}) . \quad (10.43)$$

Again, this is fine for $\lambda = 1, 2, 3$, but it is negative for $\lambda = \lambda' = 0$. The inner product therefore fails to be positive definite. This same problem can arise for multi-particle states as well.

iv) Energies

The Hamiltonian of Eq. (10.41) seems to have the wrong sign for the $a_{k,0}^\dagger a_{k,0}$ term. Applying it to the one-particle state $|p, \lambda = 0\rangle$, one finds

$$H|p, 0\rangle = -p^0|p, 0\rangle , \quad (10.44)$$

so the eigenvalue is negative. However, the expectation value of the Hamiltonian for this state is positive due to the negative inner product of the state with itself,

$$\langle p, 0|H|p, 0\rangle = -p^0\langle p, 0|p, 0\rangle = +p^0|\langle p, 0|p, 0\rangle| . \quad (10.45)$$

v) Lorentz Gauge

Things look bad, but we will see that gauge invariance saves things. However, implementing gauge is a little bit tricky. For example, just forcing the Lorenz gauge condition $\partial_\mu A^\mu = 0$ is not consistent with the ETCRs. Explicitly,

$$[\partial_\mu A^\mu(t, \vec{x}), A^\nu(t, \vec{x}')] = i\eta^{\nu 0}\delta^{(3)}(\vec{x} - \vec{x}') \neq 0 . \quad (10.46)$$

10.2.4 Imposing the Gauge Condition

Recall that we started off with the free QED Lagrangian, added the *gauge fixing* term $-(\partial_\mu A^\mu)/2\xi$ to it as in Eq. (10.27), and then set $\xi = 1$ (rather than treating it as a genuine Lagrange multiplier). With this modified Lagrangian, we derived the conjugate momenta and the classical Hamiltonian, and we applied the usual canonical quantization procedure to it. The resulting quantum theory seems to have many undesirable properties.

The way to fix up the theory to describe physical photons is to find a way to impose the gauge condition $\partial_\mu A^\mu = 0$. Just imposing this condition as an operator equation is too strong, because it conflicts with the ETCRs. Instead, we will only demand that this condition hold in a weakened form when acting on a subset of states in the Hilbert that we will identify with physical configurations of the system. This method is sometimes called the Gupta-Bleuler method after the two people who came up with it.

The key insight of Gupta and Bleuler is that only a subset of the states in the Hilbert should be identified with physical particle excitations. Given a state $|\psi\rangle$, the “physicality” condition is

$$\partial_\mu A_-^\mu |\psi\rangle = 0 , \quad (10.47)$$

where

$$A_-^\mu(x) = \sum_{\lambda=0}^4 \int \widetilde{dk} a_{k,\lambda} \epsilon^\mu(k, \lambda) e^{-ik \cdot x} . \quad (10.48)$$

Note that $A^\mu = A_-^\mu + A_+^\mu$ with $A_+^\mu = (A_-^\mu)^\dagger$. Applying the derivative to this expansion and using the properties of the polarization 4-vectors, Eq. (10.47) implies that

$$L_k |\psi\rangle := (a_{k,0} - a_{k,3}) |\psi\rangle = 0 . \quad (10.49)$$

Taking the conjugate, we also have $\langle 0 | L_k^\dagger = 0$.

The G-B condition is trivially satisfied by the vacuum $|0\rangle$. Going to 1-particle states, $|k, 1\rangle$ and $|k, 2\rangle$ are both physical excitations, but $|k, 0\rangle$ and $|k, 3\rangle$ individually are not. The standard terminology is that the $\lambda = 1, 2$ are called the *transverse* modes, $\lambda = 3$ is *longitudinal*, and $\lambda = 0$ is *timelike*. More generally, we have for any physical state $|\psi\rangle$

$$\langle \psi | a_{k,3}^\dagger a_{k,3} |\psi\rangle = \langle \psi | a_{k,0}^\dagger a_{k,0} |\psi\rangle , \quad (10.50)$$

as well as

$$\langle \psi | H |\psi\rangle = \sum_{\lambda=1}^2 \int \widetilde{dk} k^0 a_{k,\lambda}^\dagger a_{k,\lambda} . \quad (10.51)$$

Thus, only the transverse modes contribute to the (non-negative) energy of a physical state.

Things are looking up, but there is still the question of how to interpret the longitudinal and timelike polarizations. While neither $|k, 0\rangle$ nor $|k, 3\rangle$ are physical, there is a particular linear combination of them that is,

$$L_k^\dagger |0\rangle = |k, 0\rangle - |k, 3\rangle . \quad (10.52)$$

The physicality of this state follows from $[L_k, L_k^\dagger] = 0$. This state also has $HL_k^\dagger |0\rangle = 0$. We interpret this state as being physically equivalent to the vacuum $|0\rangle$ (up to a possible

normalization factor). More generally, it can be shown that any physical state can be written in the form [19]

$$|\psi\rangle = G|\psi_T\rangle, \quad (10.53)$$

where $|\psi_T\rangle$ contains only transverse photons (with the same transverse content as $|\psi\rangle$) and the operator G consists exclusively of sums and products of L_k^\dagger operators.³ Furthermore, one can also show that [19]

$$\langle\psi|A^\mu(x)|\psi\rangle = \langle\psi_T|A^\mu(x)|\psi_T\rangle + \langle\psi_T|\partial^\mu\alpha(x)|\psi_T\rangle, \quad (10.54)$$

for some scalar function α . Thus, we also interpret $|\psi\rangle$ and the corresponding $|\psi_T\rangle$ states as being physically equivalent, and related by a gauge transformation.

10.2.5 Propagation

To derive Feynman rules, we will need to generalize Wick's theorem to include photon field contractions. As before, the contraction is equal to the time-ordered 2-point function. Applying the mode expansion of $A^\mu(x)$ and the completeness relations for the polarization 4-vectors, the result is

$$\langle 0|T\{A^\mu(x)A^\nu(x')\}|0\rangle = \int \frac{d^4k}{(2\pi)^4} \frac{-i\eta^{\mu\nu}}{k^2 + i\epsilon} e^{-ik\cdot(x-x')}. \quad (10.55)$$

We call this the photon propagator.

Two comments about this result are in order. First, all four polarizations contribute to it, which should not be too surprising given that the operator $A^\mu A^\nu$ is not gauge-invariant. Second, this result corresponds to the specific choice of $\xi = 1$ for the gauge fixing parameter. Going through the same procedure for an arbitrary value of ξ , the propagator turns out to be

$$\langle 0|T\{A^\mu(x)A^\nu(x')\}|0\rangle = \int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2 + i\epsilon} \left[-\eta^{\mu\nu} + (1 - \xi) \frac{k^\mu k^\nu}{k^2} \right] e^{-ik\cdot(x-x')}. \quad (10.56)$$

Our choice of $\xi = 1$ is sometimes called Feynman gauge. Other useful values are $\xi = 0$ (Landau gauge) and $\xi = 3$ (Yennie gauge). We will see shortly that ξ should not appear in any physical observable.

10.3 Interacting Photons

Let us turn next to QED, with its fermion interactions,

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\Psi}i\gamma^\mu D_\mu\Psi - m\bar{\Psi}\Psi \quad (10.57)$$

$$= (\text{free theory}) - A_\mu(eQ\bar{\Psi}\gamma^\mu\Psi). \quad (10.58)$$

³More precisely, $G = 1 + \int \widetilde{d}k_1 A_1(k_1)L_{k_1}^\dagger + \int \widetilde{d}k_1 \int \widetilde{d}k_2 A_2(k_1, k_2)L_{k_1}^\dagger L_{k_2}^\dagger + \dots$

From this, we see that the only coupling in QED connects the vector field to $j^\mu = Q\bar{\Psi}\gamma^\mu\Psi$, the conserved Noether current corresponding to the symmetry of the theory under global rephasing. This has an important implication for scattering matrix elements computed in the theory.

Consider a general QED scattering matrix element with an external photon of momentum k . Given our Feynman rules, the matrix element will have the form (for an incoming photon)

$$\mathcal{M} = \epsilon^\mu(k, \lambda) \mathcal{M}_\mu(k) , \quad (10.59)$$

for some quantity \mathcal{M}_μ . Since A_μ always connects with a j^μ in a vertex, we have $\mathcal{M}^\mu(k)$ [1]

$$\mathcal{M}_\mu(k) \sim LSZ \int d^4x e^{-ik\cdot x} \langle j_\mu(x) \mathcal{O} \rangle , \quad (10.60)$$

where \mathcal{O} is whatever operator that is needed to make up the rest of the amplitude. This implies that

$$k_\mu \mathcal{M}^\mu(k) = i \int d^4x (\partial_\mu e^{-ik\cdot x}) \langle j^\mu(x) \mathcal{O} \rangle \quad (10.61)$$

$$= -i \int d^4x e^{-ik\cdot x} \langle \partial_\mu j^\mu(x) \mathcal{O} \rangle \quad (10.62)$$

$$= 0 , \quad (10.63)$$

where we have integrated by parts to get the second line. This result, $k_\mu \mathcal{M}^\mu(k)$ is called the *Ward identity*. It is on account of the Ward identity that we can ignore any p_μ terms in the photon polarization sums, so that $\sum_\lambda \epsilon^\mu(p, \lambda) \epsilon^{\nu*}(p, \lambda) \rightarrow -\eta^{\mu\nu}$.

Chapter 11

Path Integral Quantization

In formulating quantum theories based on classical field theories, we have used the so-called *canonical quantization* procedure, where we elevated the classical Poisson brackets to (anti-)commutators between operators on a Hilbert space. It turns out there exists a second, equivalent way to formulate quantum theories called the *path integral* method. In some cases, and for theories with gauge invariance in particular, path integrals are much more convenient than canonical quantization. Understanding the path integral method also provides new insights on what we have already done.

You may or may not have encountered path integrals in your previous quantum mechanics courses. If not, don't worry since we will present everything you need to know here. Also, don't worry too much about mathematical exactitude. We don't really know how to define path integrals in a fully rigorous way. Nevertheless, they are extremely useful and they give a very nice intuitive picture of quantum mechanics. My favourite treatment of path integrals is Appendix A of Ref. [20], and much of these notes follows this text.¹

11.1 Path Integrals for Scalar Fields

Let's begin with the easiest case, namely the scalar field $\phi(x)$. We will begin by defining path integrals as general mathematical objects. Next, we will show how they can be used to formulate a quantum theory of scalar fields.

11.1.1 Introduction to Path Integrals

Recall from `notes-02` that we defined the functional derivative with respect to a scalar field $\phi(x)$ according to

$$\frac{\delta f(\phi(x))}{\delta \phi(x')} = \frac{\partial f}{\partial \phi} \delta^{(4)}(x - x') . \quad (11.1)$$

¹You don't need to know any string theory to understand the appendix.

This is just a continuous generalization of $\partial_\mu f(x) = \delta_\mu^\nu (\partial f / \partial x^\nu)$. Functional (or path) integrals will turn out to be an analogous generalization of regular integrals.

Let us first define the functional integral of the field $\phi(t, \vec{x})$ at the fixed time t_i . For this, we divide space into a lattice of points $\{\vec{x}_j\}$ labelled by $j = 1, \dots, M$. The functional integral at time t_i is then defined to be

$$\int [\mathcal{D}'\phi_i] \sim \lim_{M \rightarrow \infty} \prod_{j=1}^M \left[\int_{-\infty}^{\infty} d\phi(t_i, \vec{x}_j) \right], \quad (11.2)$$

where the squiggle means equality up to an overall factor. The functional integral is therefore just the product of integrals over the field values at each point in space. As you can imagine, taking the *continuum limit* $M \rightarrow \infty$ is highly non-trivial, and not even necessarily well-defined.

To define the full path integral, let us take the time interval $[t', t'']$ and subdivide it into $(N+1)$ pieces. Take $t' = t_0$ and $t'' = t_{N+1}$, with t_i as the i -th intermediate time slice. The full path integral with fixed endpoints $\phi(t', \vec{x}) = \phi'(\vec{x})$ and $\phi(t'', \vec{x}) = \phi''(\vec{x})$ is then

$$\int [\mathcal{D}\phi]_{\phi'}^{\phi''} \sim \lim_{N \rightarrow \infty} \prod_{i=1}^N \int [\mathcal{D}'\phi_i]. \quad (11.3)$$

Note that the endpoints are fixed, and do not get integrated over. In many cases we will take $t' \rightarrow -\infty$, $t'' \rightarrow \infty$, and force $\phi', \phi'' \rightarrow 0$. We will write the resulting path integral in this case as simply $\int [\mathcal{D}\phi]$. It is equivalent to the more symmetric expression

$$\int [\mathcal{D}\phi] \sim \prod_x \left[\int_{-\infty}^{\infty} d\phi(t, \vec{x}) \right]. \quad (11.4)$$

You should think of this as a sum over all possible spacetime configurations of the field ϕ . The product over points in spacetime $x = (t, \vec{x})$ can also be viewed as the product over a lattice of points $x_i = (t_i, \vec{x}_i)$ in the limit that the lattice spacing is taken to zero.

The bad news about path integrals is that this definition is rather less than precise. However, the good news is that we only ever do a small number of different types of integrals. The first, and easiest, is

$$\int [\mathcal{D}\phi] \delta[\phi - \phi'] S[\phi] = S[\phi'], \quad (11.5)$$

where $S[\phi]$ is any functional of the fields and ϕ' is a specific function (*i.e.* a specific field configuration). The delta functional can also be written as a path integral, just like we have for the usual delta function. To derive it, note first that the delta functional is the product of delta functions at each point in spacetime, since both are zero unless the argument is the zero function:

$$\delta[\phi] \sim \prod_x \delta(\phi_x), \quad (11.6)$$

where $\phi_x = \phi(x)$ is the value of the field at point x (and not a function of x). We also have

$$\delta(\phi_x) = \int \frac{d\omega_x}{2\pi} e^{i\omega_x \phi_x} . \quad (11.7)$$

It follows that

$$\delta[\phi] \sim \prod_x \left(\int d\omega_x e^{i\omega_x \phi_x} \right) \quad (11.8)$$

$$\sim \left(\prod_x \int d\omega_x \right) \exp\left(i \sum_x \omega_x \phi_x\right) \quad (11.9)$$

$$\sim \int [\mathcal{D}\omega] \exp \left[i \int d^4x \omega(x) \phi(x) \right] . \quad (11.10)$$

The second kind of integral we will encounter is the Gaussian. Recall that

$$\int_{-\infty}^{\infty} dx e^{-\alpha x^2} = \sqrt{\frac{\pi}{\alpha}} . \quad (11.11)$$

The path integral Gaussian is just a multi-dimensional generalization of this. To compute it, recall that the set of all functions on spacetime is an infinite-dimensional vector space with an inner product given by

$$f \cdot g = \int d^4x f(x)g(x) . \quad (11.12)$$

An element of the space is just a function f , and $f(x)$ is the x -th component of the function.²

The typical integrand of a functional Gaussian integral is

$$\exp[-\phi \cdot (\Delta\phi)] = \exp \left[- \int d^4x \phi(x) \Delta \phi(x) \right] , \quad (11.13)$$

where Δ is a differential operator with only even powers of ∂^2 . Since we assume that ϕ vanishes at the boundary, this operator is Hermitian with respect to the inner product. It follows that we can expand any function ϕ in terms of a set of basis functions $\{f_A\}$ such that

$$\Delta f_A = \lambda_A f_A , \quad (11.14)$$

and

$$\int d^4x f_A f_B = \delta_{AB} . \quad (11.15)$$

Expanding ϕ in terms of these basis functions gives

$$\phi(x) = \sum_A \phi_A f_A(x) , \quad (11.16)$$

²You've seen this in QM: if $|\psi\rangle$ is a state, $\psi(x) = \langle x|\psi\rangle$ is the x -th component in the position basis.

where ϕ_A is the A -th expansion coefficient. The Gaussian integrand then becomes

$$\exp[-\phi \cdot (\Delta\phi)] = \exp\left[-\sum_A \lambda_A \phi_A^2\right] = \prod_A e^{-\lambda_A \phi_A^2} . \quad (11.17)$$

Since any field configuration can be specified completely by the expansion coefficients ϕ_A , integrating over them is equivalent to doing the path integral. Thus, we have

$$\int [\mathcal{D}\phi] \sim \prod_A \left[\int_{-\infty}^{\infty} d\phi_A \right] . \quad (11.18)$$

Putting these pieces together, we see that

$$\int [\mathcal{D}\phi] e^{-\phi \cdot (\Delta\phi)} \sim \prod_A \left[\int d\phi_A e^{-\lambda_A \phi_A^2} \right] \quad (11.19)$$

$$\sim \prod_A \left(\frac{\pi}{\lambda_A} \right)^{1/2} \quad (11.20)$$

$$\sim (\det \Delta)^{-1/2} . \quad (11.21)$$

Note that defining the determinant of an operator to be the product of its eigenvalues coincides with what we would find for a finite-dimensional diagonalizable matrix.

11.1.2 Path Integral Quantization of the Free Scalar

In [notes-02](#), we discussed eigenstates of the Schrödinger-picture field operator at the fixed reference time $t = 0$, $\hat{\phi}(\vec{x}) = \hat{\phi}(0, \vec{x})$:³

$$\hat{\phi}(\vec{x})|\phi'\rangle = \phi'(\vec{x})|\phi'\rangle . \quad (11.22)$$

On the left-hand side we have the field operator (denoted by the hat), while on the right-hand side we have the specific classical field function that is its eigenvalue on the state $|\phi'\rangle$. In the Heisenberg picture, we have

$$\hat{\phi}(x) = \hat{\phi}(t, \vec{x}) = e^{iHt} \hat{\phi}(0, \vec{x}) e^{-iHt} . \quad (11.23)$$

In this picture, we can also form comoving eigenstates defined by

$$|\phi'(t)\rangle = e^{iHt} |\phi'\rangle . \quad (11.24)$$

These obviously satisfy

$$\hat{\phi}(t, \vec{x}) |\phi'(t)\rangle = \phi'(\vec{x}) |\phi'(t)\rangle . \quad (11.25)$$

Comoving eigenstates evolve oppositely to physical states in the Schrödinger picture.

³In these notes, we will use a hat to denote a quantum operator.

The key result of path integral quantization is [20]⁴

$$\langle \phi''(t'') | T \{ \hat{\phi}(x_1) \dots \hat{\phi}(x_n) \} | \phi'(t') \rangle = \int [\mathcal{D}\phi]_{\phi'}^{\phi''} \exp \left[i \int_{t'}^{t''} dt L(\phi, \dot{\phi}) \right] \phi(x_1) \dots \phi(x_n) , \quad (11.26)$$

where $t' < t_1, \dots, t_n < t''$. This is a very deep relation. On the left side is a quantum mechanical matrix element of an operator. On the right, we just have a complicated sum over *classical* field configurations weighted by the action!

For scattering calculations, we only need to find the expectation values of operators in the vacuum state $|\Omega\rangle$ of the theory. Similarly to what we did in **notes-03**, we can project it out by taking the limit of $t \rightarrow \pm\infty$ in a slightly imaginary direction. For example,

$$\lim_{t \rightarrow -\infty(1-i\epsilon)} |\phi'(t)\rangle = \lim_{t \rightarrow -\infty(1-i\epsilon)} e^{iHt} \left(\sum_{\lambda} |\lambda\rangle \langle \lambda| \right) |\phi'\rangle \quad (11.27)$$

$$= \lim_{t \rightarrow -\infty(1-i\epsilon)} \left[|\Omega\rangle \langle \Omega| \phi'\rangle + \sum_{\lambda \neq \Omega} e^{iE_{\lambda}t} |\lambda\rangle \langle \lambda| \phi'\rangle \right] \quad (11.28)$$

$$= (\langle \Omega | \phi'\rangle) |\Omega\rangle \quad (11.29)$$

Instead of tilting t , we can instead add a small perturbation to the energy to make it slightly imaginary. A convenient way to do this in a field theory is to make the mass slightly imaginary, since the corresponding ϕ^2 operator is positive definite,

$$-\mathcal{L} \supset m^2 \phi^2 \rightarrow (m^2 - i\epsilon) \phi^2 , \quad (11.30)$$

with $\epsilon > 0$. This induces

$$\mathcal{H}(m^2 - i\epsilon) \rightarrow \mathcal{H}(m^2) + \frac{\partial \mathcal{H}}{\partial m^2} (-i\epsilon) = \mathcal{H}(m^2) - i\epsilon' , \quad (11.31)$$

with $\epsilon' > 0$ since $\partial \mathcal{H} / \partial m^2 > 0$ in any sensible theory. This gives

$$E_{\lambda} \rightarrow E_{\lambda} - i\epsilon' , \quad (11.32)$$

which projects out the vacuum when inserted in Eq. (11.28).

After applying this projection, we obtain a master formula for path integrals:

$$\langle \Omega | T \{ \hat{\phi}(x_1) \dots \hat{\phi}(x_n) \} | \Omega \rangle = \int [\mathcal{D}\phi] e^{iS'[\phi]} \phi(x_1) \dots \phi(x_n) / \int [\mathcal{D}\phi] e^{iS'[\phi]} , \quad (11.33)$$

where the S' is the action with a slightly imaginary mass, and we have normalized the right-hand side so that $\langle 1 \rangle = 1$. This formula should remind you of a statistical mechanical partition function, with $e^{iS'}$ instead of $e^{-H/T}$ as the exponential weight.

⁴See Refs. [20, 3, 1, 2] for proofs.

To study operator expectation values, let us define the *generating functional* $Z[J]$ by

$$Z[J] = \int[\mathcal{D}\phi] \exp(iS[\phi] + iJ \cdot \phi) = \int[\mathcal{D}\phi] \exp \left[i \int d^4x (\mathcal{L} + J\phi) \right] , \quad (11.34)$$

where $J(x)$ is an unspecified function and $J \cdot \phi = \int d^4x J(x)\phi(x)$ as before. The function $J(x)$ is often called the source. Suppose we take the functional integral of $Z[J]$ with respect to $J(x)$. This has the effect of adding a power of $i\phi(x)$ to the integrand:

$$\frac{\delta Z[J]}{\delta J(x)} = \int[\mathcal{D}\phi] \frac{\delta}{\delta J(x)} e^{i(S[\phi] + J \cdot \phi)} \quad (11.35)$$

$$= \int[\mathcal{D}\phi] \left[\int d^4y i\phi(y) \delta^{(4)}(y - x) \right] e^{i(S[\phi] + J \cdot \phi)} \quad (11.36)$$

$$= \int[\mathcal{D}\phi] i\phi(x) e^{i(S[\phi] + J \cdot \phi)} \quad (11.37)$$

Taking more derivatives would add more powers of ϕ to the integrand. We can use this to rewrite our master formula, Eq. (11.33), as

$$\langle \Omega | T \{ \hat{\phi}(x_1) \dots \hat{\phi}(x_n) \} | \Omega \rangle = (-i)^n \frac{\delta}{\delta J(x_1)} \dots \frac{\delta}{\delta J(x_n)} Z[J] \Big|_{J=0} / Z[0] . \quad (11.38)$$

Setting $J \rightarrow 0$ after taking the derivatives gets rid of the $J \cdot \phi$ in the exponential, and normalizing by $1/Z[0]$ ensures that $\langle 1 \rangle = 1$.

So far, all our results are completely general and apply to both free and interacting scalar theories. Let us now apply them specifically to the free scalar theory,

$$\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}(m^2 - i\epsilon)\phi^2 . \quad (11.39)$$

after integrating by parts, this gives

$$S[\phi] = \frac{1}{2} \int d^4x \phi(-\partial^2 - m^2 + i\epsilon)\phi \quad (11.40)$$

$$:= \frac{1}{2} \phi \cdot (\Delta\phi) , \quad (11.41)$$

with $\Delta = (-\partial^2 - m^2 + i\epsilon)$. When inserted into the path integral, this will just be a Gaussian.

We can evaluate the generating functional for the free theory by completing the square and computing the Gaussian path integral. For this, note that

$$\Delta D_F(x) = (-\partial^2 - m^2 + i\epsilon) \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2 + i\epsilon} e^{-ip \cdot x} \quad (11.42)$$

$$= i\delta^{(4)}(x) . \quad (11.43)$$

We can therefore think of $-iD_F = \Delta^{-1}$. With this result in hand, we can complete the square by making the change of variable

$$\phi(x) = \phi'(x) - \int d^4y (-i)D_F(x - y)J(y) := \phi'(x) - (\Delta^{-1}J)(x) . \quad (11.44)$$

In the exponential of $Z[J]$, this gives

$$S[\phi] + J \cdot \phi = \frac{1}{2}(\phi' - \Delta^{-1}J) \cdot [\Delta(\phi' - \Delta^{-1}J)] + J \cdot (\phi' - \Delta^{-1}J) \quad (11.45)$$

$$= \frac{1}{2}\phi' \cdot (\Delta\phi') - \frac{1}{2}J \cdot (\Delta^{-1}J) \quad (11.46)$$

$$= \frac{1}{2} \int d^4x \phi' \Delta \phi' + \frac{i}{2} \int d^4x \int d^4y J(x) D_F(x-y) J(y) . \quad (11.47)$$

In going from the first to the second line we have implicitly integrated by parts to move the Δ operator around. The integrand in terms of ϕ' is now a Gaussian. Since the change of variables was just a shift by a constant, the path integral measure is also unchanged,

$$[\mathcal{D}\phi] = [\mathcal{D}\phi'] . \quad (11.48)$$

It follows that the free-theory generating functional is equal to

$$Z[J] \sim (\det \Delta)^{-1/2} \exp \left[-\frac{1}{2} \int d^4x \int d^4y J(x) D_F(x-y) J(y) \right] . \quad (11.49)$$

Note that the $i\epsilon$ factor inserted to project out the ground state leads to the Feynman propagator in this expression. It also ensures the convergence of the Gaussian integral.

We can now use Eq. (11.38) to compute n -point functions in the free theory. The 2-point functions comes out to be

$$(-i)^2 \frac{\delta^2}{\delta J(x) \delta J(y)} \left(e^{-\frac{1}{2} J \cdot D_F \cdot J} \right) \Big|_{J=0} = (-1)^2 D_F(x-y) , \quad (11.50)$$

which is precisely what we found previously. We also see that the propagator is just the inverse of the quadratic operator in the Lagrangian. Repeating with more derivatives, we reproduce all the n -point functions of the theory (and Wick's theorem) in a very trivial way. In this case, the contraction of two fields corresponds to the corresponding derivatives hitting the same term in the expansion of the exponential.

11.2 Path Integrals for Fermions

We can also quantize fermion fields using path integrals. For this, we need to define regular derivatives and integrals of Grassmann numbers. From here, it is straightforward to generalize to fermionic path integrals, and then to quantization.

11.2.1 Grassman Integrals, Regular and Path

Suppose we have a single Grassmann number η . Since $\eta^2 = 0$, the most general real function of η is

$$f(\eta) = a + b\eta , \quad (11.51)$$

for some real coefficients a and b . We define the derivative by

$$\frac{d}{d\eta} 1 = 0, \quad \frac{d}{d\eta} \eta = 1, \quad (11.52)$$

so that

$$\frac{d}{d\eta} f(\eta) = b. \quad (11.53)$$

For integrals, we define

$$\int d\eta 1 = 0, \quad \int d\eta \eta = 1. \quad (11.54)$$

It is best to think of $\int d\eta$ as an operator on functions of η rather than an integral in the usual sense. This definition implies that

$$\int d\eta \frac{df(\eta)}{d\eta} = 0. \quad (11.55)$$

When there are multiple Grassmann numbers, we treat the integral and the derivative as operators acting from the left, and put everything together by anticommuting. For example, given η and χ we have

$$\frac{d}{d\eta} \chi = 0, \quad \frac{d}{d\eta} (\chi\eta) = -\frac{d}{d\eta} (\eta\chi) = -\chi. \quad (11.56)$$

Integrals work in the same way.

Having warmed up with regular derivatives and integrals, let's generalize them to their functional counterparts. We have

$$\frac{\delta\Psi_a(x)}{\delta\Psi_b(y)} = \delta_a^b \delta^{(4)}(x-y). \quad (11.57)$$

For the functional integral, we take

$$\int [\mathcal{D}\Psi] \sim \prod_x \left[\int d\Psi_1(x) \int d\Psi_2(x) \int d\Psi_3(x) \int d\Psi_4(x) \right]. \quad (11.58)$$

In other words, do the Grassmann integral over each component of Ψ at every point in spacetime and multiply together all the results. Note that $\Psi_a(x)$ should be treated as independent Grassmann variables in that $\Psi_a(x)\Psi_b(x')$ is non-zero unless $a = b$ and $x = x'$.

Evaluating fermionic functional integrals is even easier than for bosons. The case of most interest to us will be the Gaussian,

$$\int [\mathcal{D}\Psi \mathcal{D}\bar{\Psi}] \exp \left[- \int d^4x \bar{\Psi} \Delta \Psi \right], \quad (11.59)$$

where Δ is a Hermitian differential operator. As before, let us expand Ψ and $\bar{\Psi}$ in orthonormal eigenfunctions of the operator Δ :

$$\Psi(x) = \sum_A \psi_A g_A(x), \quad \bar{\Psi}(x) = \sum_B \bar{\psi}_B h_B(x), \quad (11.60)$$

where g_A and h_A are bosonic functions and ψ_A and $\bar{\psi}_B$ are Grassmannian expansion coefficients. For the eigenfunctions, we have

$$\Delta g_A = \lambda_A g_A \quad \Delta h_B = \lambda_B h_B, \quad (11.61)$$

as well as

$$\int d^A x h_B(x) g_A(x) = \delta_{AB}. \quad (11.62)$$

Putting this into the integrand, we find

$$\exp[\bar{\Psi} \cdot (\Delta \Psi)] = \exp\left[-\sum_A \bar{\psi}_A \lambda_A \psi_A\right] = \prod_A e^{-\bar{\psi}_A \lambda_A \psi_A}. \quad (11.63)$$

The path integral measure becomes

$$\int[\mathcal{D}\Psi \mathcal{D}\bar{\Psi}] \sim \prod_A \left[\int d\psi_A \int d\bar{\psi}_A \right]. \quad (11.64)$$

The Gaussian integral is therefore equal to

$$\int[\mathcal{D}\Psi \mathcal{D}\bar{\Psi}] e^{-\bar{\Psi} \cdot (\Delta \Psi)} \sim \prod_A \left[\int d\psi_A \int d\bar{\psi}_A e^{-\bar{\psi}_A \lambda_A \psi_A} \right] \quad (11.65)$$

$$\sim \prod_A \left[\int d\psi_A \int d\bar{\psi}_A (1 - \lambda_A \bar{\psi}_A \psi_A) \right] \quad (11.66)$$

$$\sim \prod_A (-\lambda_A) \quad (11.67)$$

$$\sim \det(\Delta). \quad (11.68)$$

Note that the determinant has a positive power rather than a negative one. This is characteristic of fermions. The factor of +1 relative to 1/2 comes from the fact that two different fields are now involved.

11.2.2 Path Integral Quantization

With these tools in hand, we can now turn to physics. The master formula for a Dirac fermion is

$$\langle \Omega | T \{ \hat{\Psi}_{a_1}(x_1) \dots \hat{\Psi}^{b_n}(x_n) \} | \Omega \rangle = \int[\mathcal{D}\Psi \mathcal{D}\bar{\Psi}] e^{iS'} \Psi_{a_1}(x_1) \dots \bar{\Psi}^{b_n}(x_n) / \int[\mathcal{D}\Psi \mathcal{D}\bar{\Psi}] e^{iS'} \quad (11.69)$$

where $S' = S[\Psi, \bar{\Psi}] + i\epsilon\bar{\Psi}\Psi$. When there are multiple types of fields, we should integrate over each one of them in the path integral.

To compute n -point functions, it is again convenient to define a generating functional, this time with a source for each independent field. For a theory of Dirac fermions,

$$Z[\eta, \bar{\eta}] = \int [\mathcal{D}\Psi \mathcal{D}\bar{\Psi}] \exp [iS[\Psi, \bar{\Psi}] + i\bar{\eta} \cdot \Psi + i\bar{\Psi} \cdot \eta] . \quad (11.70)$$

We now have

$$\frac{\delta Z}{\delta \bar{\eta}^a(x)} = \int [\mathcal{D}\Psi \mathcal{D}\bar{\Psi}] (i\Psi_a(x)) e^{iS[\Psi, \bar{\Psi}] + i\bar{\eta} \cdot \Psi + i\bar{\Psi} \cdot \eta} , \quad (11.71)$$

$$\frac{\delta Z}{\delta \eta_b(x)} = \int [\mathcal{D}\Psi \mathcal{D}\bar{\Psi}] (-i\bar{\Psi}^b(x)) e^{iS[\Psi, \bar{\Psi}] + i\bar{\eta} \cdot \Psi + i\bar{\Psi} \cdot \eta} . \quad (11.72)$$

Note the extra minus sign from anticommuting the fermionic functional derivative. Taking multiple derivatives and setting $\eta = \bar{\eta} = 0$, we get

$$\langle \Omega | T \{ \Psi_{a_1}(x_1) \dots \bar{\Psi}^{b_n}(x_n) \} | \Omega \rangle = \frac{1}{Z[0, 0]} \left(-i \frac{\delta}{\delta \bar{\eta}^{a_1}(x_1)} \right) \dots \left(+i \frac{\delta}{\delta \eta_{b_n}(x_n)} \right) Z[\eta, \bar{\eta}] \Big|_{\eta = \bar{\eta} = 0} \quad (11.73)$$

Let's now specialize to the free theory of a Dirac fermion, with action

$$S[\Psi, \bar{\Psi}] = \int d^4x \bar{\Psi} (i\gamma \cdot \partial - m + i\epsilon) \Psi := \bar{\Psi} \cdot (\Delta \Psi) . \quad (11.74)$$

As before, we can complete the square in the generating functional and do the resulting Gaussian integral. The appropriate changes of variables in this case are

$$\Psi(x) = \Psi'(x) - \int d^4y (-i) S_F(x-y) \eta(y) \quad (11.75)$$

$$\bar{\Psi}(x) = \bar{\Psi}'(x) - \int d^4y (-i) \bar{\eta}(y) S_F(y-x) . \quad (11.76)$$

Using the fact that

$$\Delta [S_F(x)]_a^b = i\delta_a^b \delta^{(4)}(x) , \quad (11.77)$$

we obtain

$$Z[\eta, \bar{\eta}] \sim \det(\Delta) \exp \left[- \int d^4x \int d^4y \bar{\eta}(x) S_F(x-y) \eta(y) \right] . \quad (11.78)$$

As for the scalar, it is not hard to check that this formula reproduces all the n -point functions (and Wick's theorem) of the free Dirac theory.

11.3 Quantizing the Photon

Path integrals turn out to be very convenient for quantizing gauge theories. We will show how this works for the case of the free photon field.

11.3.1 Two Puzzles

The goal is to apply our path integral formalism to the free photon, with Lagrangian

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} = -\frac{1}{2}A^\mu(-\eta^{\mu\nu}\partial^2 + \partial_\mu\partial_\nu)A^\nu, \quad (11.79)$$

where we have integrated by parts to get the second equality. Given this Lagrangian, we can construct an action and formulate a master formula for n -point functions in analogy with Eq. (11.33). The direct generalization is

$$\langle\Omega|T\{\hat{A}^{\mu_1}(x_1)\dots\hat{A}^{\mu_n}(x_n)\}|\Omega\rangle \stackrel{?}{=} \int[\mathcal{D}A^\mu] A^{\mu_1}(x_1)\dots A^{\mu_n}(x_n)e^{iS'[A]} / \int[\mathcal{D}A^\mu] e^{iS'[A]} \quad (11.80)$$

where

$$\int[\mathcal{D}A^\mu] = \int[\mathcal{D}A^0] \dots \int[\mathcal{D}A^3]. \quad (11.81)$$

It turns out that this direct generalization is not quite correct.

There are two immediate problems with Eq. (11.80). The first is that it involves an integration over many physically equivalent field configurations. In fact, given a configuration $A_\mu(x)$, there are infinitely many other configurations that give the same physics, namely

$$A'_\mu(x) = A_\mu(x) + \frac{1}{e}\partial_\mu\alpha(x), \quad (11.82)$$

for any function $\alpha(x)$. The integration in Eq. (11.80) therefore involves a huge (infinite!) overcounting.

The second problem with the attempt of Eq. (11.80) has to do with the propagator. Recall that for both the scalar and the fermion, we wrote the free theory as $\phi^{(*)} \cdot (\Delta\phi)$ for some Hermitian operator Δ . In the present case, we can do the same but with

$$\Delta \rightarrow \Delta_{\mu\nu} = \eta_{\mu\nu}\partial^2 - \partial_\mu\partial_\nu. \quad (11.83)$$

If we were to follow our previous procedure and add a source and complete the square, we would obtain an operator containing the inverse of $\Delta_{\mu\nu}$. Unfortunately, this inverse is not well-defined because $\Delta_{\mu\nu}$ has many zero eigenvalues. For example,

$$\Delta_{\mu\nu}(\partial^\nu\alpha) = 0, \quad (11.84)$$

for any function α . Connecting with the first problem, we see that these zero eigenvalues correspond to field configurations that are gauge-equivalent to the zero configuration.

11.3.2 Gauge Fixing (Faddeev-Popov)

Both problems with the direct generalization of the path integral formula for scalars (and fermions) can be solved by taking into account the physics of gauge invariance. The idea

will be to find a way to factor out the gauge redundancy and integrate over only physically distinct field configurations. It will turn out that this procedure also produces an invertible propagator factor.

As a first step, let us assume we can find a *gauge fixing* condition on the field A^μ that removes the gauge redundancy [21],

$$F(A_\mu) = 0 , \quad (11.85)$$

where the condition is to be applied at all spacetime points x . We will use the specific choice $F(A^\mu) = \partial_\mu A^\mu$ here, but the procedure presented here can be applied to other conditions. In general, a field configuration $\tilde{A}^\mu(x)$ will not satisfy the gauge fixing condition. However, in this Lorentz-gauge case there always exists a gauge-equivalent configuration A^μ for which the condition is satisfied, $A^\mu = \tilde{A}^\mu + \partial^\mu/e$. The gauge function α can be obtained by solving the equation

$$\partial^2 \alpha = e(\partial_\mu A^\mu) . \quad (11.86)$$

In this way, the gauge fixing condition fixes the gauge function $\alpha(x)$, up to transformations with $\partial^2 \alpha = 0$. A good gauge fixing condition is one that always has a solution, with that solution completely fixing the gauge.

Given the field configuration A^μ , let us denote the gauge-transformed version by

$$A_\alpha^\mu(x) = A^\mu(x) + \frac{1}{e} \partial^\mu \alpha(x) \quad (11.87)$$

Next, let us also define the functional $\Delta_{FP}[A]$ by

$$1 = \Delta_{FP}[A^\mu] \int [\mathcal{D}\alpha] \delta[F(A_\alpha^\mu)] . \quad (11.88)$$

As long as the gauge condition has a solution, the delta functional will give a non-zero result for any configuration A^μ . The functional $\Delta_{FP}[A]$ is then just the inverse of the number left over. A key property of $\Delta_{FP}[A]$ is that it is gauge invariant, $\Delta_{FP}[A_{\alpha'}^\mu] = \Delta_{FP}[A^\mu]$ for any gauge transformation function $\alpha'(x)$. To see this, just plug the transformed result into the transformation and use the fact that successive transformations by α' and α are equivalent to a single transformation by $\alpha'' = (\alpha + \alpha')$. Explicitly,

$$\Delta_{FP}^{-1}[A_{\alpha'}^\mu] = \int [\mathcal{D}\alpha] \delta[F(A_{\alpha'+\alpha}^\mu)] \quad (11.89)$$

$$= \int [\mathcal{D}\alpha''] \delta[F(A_{\alpha''}^\mu)] \quad (11.90)$$

$$= \Delta_{FP}^{-1}[A^\mu] . \quad (11.91)$$

In the second line we have used $[\mathcal{D}\alpha''] = [\mathcal{D}\alpha]$, since the two differ only by a shift by a fixed function $\alpha'(x)$.

Let's take this result and plug it into the naïve path integral expression we attempted to use in Eq. (11.80),

$$\int[\mathcal{D}A^\mu] \mathcal{O}(A^\mu) e^{iS[A^\mu]} = \int[\mathcal{D}A^\mu] \left(\Delta_{FP}[A^\mu] \int[\mathcal{D}\alpha] \delta[F(A^\mu_\alpha)] \right) \mathcal{O}(A^\mu) e^{iS[A^\mu]}, \quad (11.92)$$

where $\mathcal{O}(A^\mu)$ is a gauge-invariant time-ordered operator built out of A^μ fields, and we have just inserted unity in the form of Eq. (11.88). We can reorganize this expression by interchanging the orders of the functional integrals, and using gauge invariance in the form $\Delta_{FP}(A^\mu_\alpha) = \Delta_{FP}(A^\mu)$, $S[A^\mu_\alpha] = S[A^\mu]$, and $\mathcal{O}(A^\mu_\alpha) = \mathcal{O}(A^\mu)$. This gives

$$\int[\mathcal{D}A^\mu] \mathcal{O}(A^\mu) e^{iS[A^\mu]} = \int[\mathcal{D}\alpha] \int[\mathcal{D}A^\mu_\alpha] \Delta_{FP}[A^\mu_\alpha] \delta[F(A^\mu_\alpha)] \mathcal{O}(A^\mu_\alpha) e^{iS[A^\mu_\alpha]} \quad (11.93)$$

$$= \left(\int[\mathcal{D}\alpha] \right) \int[\mathcal{D}A^\mu] \Delta_{FP}[A^\mu] \delta[F(A^\mu)] \mathcal{O}(A^\mu) e^{iS[A^\mu]} \quad (11.94)$$

where in the last line we have used $[\mathcal{D}A^\mu_\alpha] = [\mathcal{D}A^\mu]$, since the integration variables only differ by a constant shift.

The result of Eq. (11.94) suggests how to build a reasonable quantum theory of the photon. In this expression, we have reorganized the simple path integral over all gauge field configurations into an expression that is independent of gauge times a functional integral over all gauge transformation parameters. This gauge factor is sometimes called the *volume* of the gauge group, and is formally infinite. It is also precisely the factor corresponding to the number of times we are overcounting physically equivalent field configurations. The sensible thing to do, therefore, is cancel it off and reinterpret the remaining factor as the correct expression for the time ordered, gauge-invariant operator matrix element:

$$\langle \Omega | T \{ \mathcal{O}(A^\mu) \} | \Omega \rangle = \int[\mathcal{D}A^\mu] \Delta_{FP}[A^\mu] \delta[F(A^\mu)] \mathcal{O}(A^\mu) e^{iS[A^\mu]} \quad (11.95)$$

$$\Big/ \int[\mathcal{D}A^\mu] \Delta_{FP}[A^\mu] \delta[F(A^\mu)] e^{iS[A^\mu]}. \quad (11.96)$$

This expression looks just like Eq. (11.80), except that we now have an additional factor of $\Delta_{FP}[A^\mu] \delta[F(A^\mu)]$ in both the numerator and denominator.

To evaluate Eq. (11.95) for a given operator, we need to figure out how to handle $\Delta_{FP}[A^\mu]$, defined by Eq. (11.88). This will require some more functional funny business. Before addressing that, let us start off with a regular multi-dimensional integral. Recall that here, changing integration variables generates a Jacobian factor. For example, consider $(x^1, x^2) \rightarrow (u^1, u^2)$, and the integral

$$\int d^2x \delta^{(2)}(u - u_0) = \int d^2u \left[\det \left(\frac{\partial u^i}{\partial x^j} \right) \right]^{-1} \delta^{(2)}(u - u_0) = \left[\det \left(\frac{\partial u^i}{\partial x^j} \right) \right]_{u=u_0}^{-1}. \quad (11.97)$$

The remaining determinant is the so-called Jacobian, and it is to be evaluated at points where $u = u_0$.

We would like to do the same type of change of variables for the functional integral relation of Eq. (11.88). First, note that $F(A_\alpha^\mu)$ should be set to zero at each point x , and to keep track of this, we will write $F(x, A_\alpha^\mu)$. The change of variables will therefore be $\alpha(y) \rightarrow F(x, A_\alpha^\mu)$. In direct analogy with the regular multi-dimensional integral, this change of variables produces a Jacobian factor

$$\Delta_{FP}^{-1}(A^\mu) = \int [\mathcal{D}\alpha] \delta[F(x, A_\alpha^\mu)] = \left| \det \left[\frac{\delta F(x, A_\alpha^\mu)}{\delta \alpha(y)} \right] \right|_{F(A_\alpha)=0}^{-1}. \quad (11.98)$$

This looks like a mess, but we can evaluate it for the specific case of $\partial_\mu A_\alpha^\mu(x) = 0$. Using gauge invariance, we can choose our representative A^μ such that it satisfies $F(x, A^\mu) = 0$ with $\alpha(x) = 0$. This implies that

$$\Delta_{FP}(A^\mu) = \left| \det \left[\frac{\delta F(x, A_\alpha^\mu)}{\delta \alpha(y)} \right] \right| = \left| \det \left(\frac{1}{e} \partial_x^2 \delta^{(4)}(x-y) \right) \right|. \quad (11.99)$$

Note that this is independent of A^μ . We can rewrite this determinant in terms of a fermionic functional integral,

$$\Delta_{FP} \sim \int [\mathcal{D}c][\mathcal{D}\bar{c}] \exp \left[\frac{i}{e} \int d^4x \bar{c} \partial^2 c \right]. \quad (11.100)$$

The fields c and \bar{c} are fermions called *Faddeev-Popov ghosts*. Since they don't couple to anything and don't appear as physical external particles, we can ignore them for QED. However, in more complicated gauge theories they couple to the gauge field and must be included in calculations at loop level.

As a final trick, let us rewrite the $\delta[F(A^\mu)]$ delta functionals in a nicer form. To do so, we modify the gauge condition to $F(A_\alpha^\mu) - \omega(x) = 0$ for some fixed function ω . All our previous work goes through in exactly the same way, with the only change being the replacement $\delta[F(A^\mu)] \rightarrow \delta[F(A^\mu) - \omega]$. Next, let us sum over many different functions $\omega(x)$ and multiply by a weighting functional $G[\omega]$ of the form

$$G[\omega] = \exp \left[\frac{-i}{2\xi} \int d^4x \omega^2(x) \right]. \quad (11.101)$$

At the end of the day, this gives (in the numerator of Eq. (11.95))

$$\int [\mathcal{D}A^\mu] \int [\mathcal{D}\omega] \delta[F(A^\mu) - \omega] G[\omega] \Delta_{FP} \mathcal{O}(A^\mu) e^{iS[A^\mu]} \quad (11.102)$$

$$= \int [\mathcal{D}A^\mu] \int [\mathcal{D}\omega] G[F(A^\mu)] \Delta_{FP} \mathcal{O}(A^\mu) e^{iS[A^\mu]} \quad (11.103)$$

$$= \int [\mathcal{D}A^\mu] \int [\mathcal{D}c][\mathcal{D}\bar{c}] \mathcal{O}(A^\mu) e^{i(S+S_{gf}+S_c)}, \quad (11.104)$$

where

$$S_{tot} = S + S_{gf} + S_c = \int d^4x \left[-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2\xi} (\partial_\mu A^\mu)^2 + \frac{1}{e} \bar{c} \partial^2 c \right]. \quad (11.105)$$

Note that we have used the delta functional to remove the integration over ω and replace it everywhere by $F(A^\mu)$.

We have just done a lot of functional heavy lifting, but the final result is very simple. The procedure of cancelling off the gauge redundancy of the functional integral amounts to adding fermionic ghost fields and some new terms in the action. In summary, removing the gauge redundancy of the functional integral can be reduced simply modifying the naïve formula of Eq. (11.80) by:

1. Include Dirac fermion ghost fields c and \bar{c} with action S_c . The fields do not couple to anything in QED and can be neglected. (This is not true in more complicated theories like QCD!)
2. Add a *gauge-fixing term* $-(\partial_\mu A^\mu)^2/2\xi$ to the action. This is the same form as the extra term we added in Gupta-Bleuler quantization.

That's it!

11.3.3 Propagation

Our complicated gauge-fixing procedure has removed the overcounting of configurations in the functional integral. Let us next address the second problem of a non-invertible photon quadratic term. It turns out that gauge fixing solves this problem too by adding a new piece to the action. After integrating by parts, we now have

$$S_{tot} = \int d^4x \frac{1}{2} A^\mu [\eta_{\mu\nu} \partial^2 - (1 - 1/\xi) \partial_\mu \partial_\nu] A^\nu . \quad (11.106)$$

This is invertible for any finite ξ . Introducing a source and completing the square as before, we obtain

$$\langle \Omega | T \{ A^\mu(x) A^\nu(x') \} | \Omega \rangle = \int d^4k e^{-ik \cdot (x-x')} \frac{i}{k^2 + i\epsilon} \left[-\eta^{\mu\nu} + (1 - \xi) \frac{k^\mu k^\nu}{k^2} \right] . \quad (11.107)$$

Again, this matches what we had in notes-10.

11.4 Interacting Theories

We have already used path integrals to reproduce all the operator expectation values of the free scalar, fermion, and photon theories studied previously. In this section we show how to use path integrals to compute operator expectation values in interacting theories.

To start, consider the interacting scalar theory given by

$$\mathcal{L} = \frac{1}{2} (\partial\phi)^2 - \frac{1}{2} m^2 \phi^2 - \frac{\lambda}{4!} \phi^4 \quad (11.108)$$

$$= \mathcal{L}_0 - \Delta V(\phi) . \quad (11.109)$$

Let us also write the generating functional of the free theory as $Z_0[J]$,

$$Z_0[J] := \int [\mathcal{D}\phi] e^{i(S_0 + J \cdot \phi)} \sim e^{-\frac{1}{2} J \cdot D_F \cdot J} , \quad (11.110)$$

where $S_0 = \int d^4x \mathcal{L}_0$. This is a nice result, but what we really want is the full generating functional $Z[J]$ so that we can compute n -point functions using Eq. (11.38).

In general, we do not know how to compute the full functional $Z[J]$ exactly. However, it is possible to derive a simple relation between $Z[J]$ and $Z_0[J]$ that lends itself to perturbation theory. For this, note that

$$\int d^4x \Delta V \left(-i \frac{\delta}{\delta \phi(x)} \right) Z_0[J] = \int [\mathcal{D}\phi] \left(\int d^4x \frac{\lambda}{4!} \phi^4(x) \right) e^{i(S_0 + iJ \cdot \phi)} , \quad (11.111)$$

where $V \left(-i \frac{\delta}{\delta \phi(x)} \right)$ means to replace every appearance of $\phi(x)$ in ΔV with $\delta/\delta\phi(x)$. Thus,

$$Z[J] = \exp \left[-i \int d^4x \Delta V \left(-i \frac{\delta}{\delta \phi(x)} \right) \right] Z_0[J] . \quad (11.112)$$

This together with Eq. (11.38) are the path integral form of the master formula from **notes-03**. Computing n -point functions is now just a matter of taking lots of derivatives.

To illustrate this procedure in a more complicated example, let us compute $\langle \phi(x_1) \Psi_a(x_2) \bar{\Psi}^b(c_3) \rangle$ in QED Lite. Introducing sources for the scalar and fermion fields as in the free theory, we have

$$Z[J, \eta, \bar{\eta}] = \exp \left[-iy \int d^4z \left(-i \frac{\delta}{\delta J(z)} \right) \left(+i \frac{\delta}{\delta \eta_c(z)} \right) \left(-i \frac{\delta}{\delta \bar{\eta}^c(z)} \right) \right] Z_0[J, \eta, \bar{\eta}] , \quad (11.113)$$

where

$$Z_0[J, \eta, \bar{\eta}] = Z_0[0, 0, 0] e^{-\frac{1}{2} J \cdot D_F \cdot J} e^{-\bar{\eta} \cdot S_F \cdot \eta} . \quad (11.114)$$

To get the 3-point function, we should differentiate with respect to the corresponding sources at x_1 , x_2 , and x_3 . Thus,

$$\langle \phi(x_1) \Psi_a(x_2) \bar{\Psi}^b(c_3) \rangle = \left(-i \frac{\delta}{\delta J_1} \right) \left(i \frac{\delta}{\delta \eta_{a2}} \right) \left(-i \frac{\delta}{\delta \bar{\eta}_3^b} \right) Z[J, \eta, \bar{\eta}] \Big|_{J=0=\eta=\bar{\eta}} / Z[0, 0, 0] \quad (11.115)$$

To evaluate this, we expand the exponential in Eq. (11.113). At leading order in the coupling y , we have

$$\langle \phi(x_1) \Psi_a(x_2) \bar{\Psi}^b(c_3) \rangle \quad (11.116)$$

$$= (iy) \frac{\delta}{\delta J_1} \frac{\delta}{\delta \eta_{a2}} \frac{\delta}{\delta \bar{\eta}_3^b} \int d^4z \frac{\delta}{\delta J(z)} \frac{\delta}{\delta \eta_c(z)} \frac{\delta}{\delta \bar{\eta}^c(z)} e^{-\frac{1}{2} J \cdot D_F \cdot J} e^{-\bar{\eta} \cdot S_F \cdot \eta} \Big|_{J=0=\eta=\bar{\eta}} \quad (11.117)$$

$$= (-iy) \int d^4z D_F(x_1 - z) [S_F(x_2 - z)]_a^c [S_F(z - x_3)]_c^b + (\text{disconnected}) \quad (11.118)$$

This matches what we found in **notes-08**.

Bibliography

- [1] M. E. Peskin and D. V. Schroeder, “An Introduction To Quantum Field Theory,” *Reading, USA: Addison-Wesley (1995) 842 p*
- [2] M. Srednicki, “Quantum field theory,” *Cambridge, UK: Univ. Pr. (2007) 641 p*
- [3] L. H. Ryder, “Quantum Field Theory,” *Cambridge, Uk: Univ. Pr. (1985) 443p*
- [4] D. J. Griffiths, “Introduction to Electrodynamics”, *Saddle River, USA: Prentice Hall (1999)*
- [5] H. Goldstein, “Classical Mechanics”, *Cambridge USA: Addison-Wesley (1950)*
- [6] D. J. Griffiths, “Introduction to Quantum Mechanics”, *Saddle River, USA: Prentice Hall (2004)*
- [7] J. Mathews, R. L. Walker, “Mathematical Methods of Physics”, *Redwood City, USA: Addison Wesley (1970)*
- [8] B. Hatfield, “Quantum field theory of point particles and strings,” *Redwood City, USA: Addison-Wesley (1992) 734 p. (Frontiers in physics, 75)*
- [9] L. S. Brown, “Quantum Field Theory,” *Cambridge, UK: Univ. Pr. (1992) 547 p*
- [10] D. Griffiths, “Introduction to elementary particles,” *Weinheim, Germany: Wiley-VCH (2008) 454 p*
- [11] H. Georgi, “Lie Algebras In Particle Physics. From Isospin To Unified Theories,” *Front. Phys.* **54**, 1-255 (1982).
- [12] M. Luke, “PHY2403F Lecture Notes,”
http://www.physics.utoronto.ca/~luke/PHY2403/References_files/lecturenotes.pdf.
- [13] D. Soper, “Lecture Notes on QFT” <http://physics.uoregon.edu/~soper/QFT/>
- [14] S. Weinberg, *Cambridge, UK: Univ. Pr. (1995) 609 p*
- [15] See Appendix A in:
H. E. Haber, G. L. Kane, “The Search for Supersymmetry: Probing Physics Beyond the Standard Model,” *Phys. Rept.* **117**, 75-263 (1985).

- [16] See Appendices A and B in :
J. Wess, J. Bagger, “Supersymmetry and supergravity,” Princeton, USA: Univ. Pr. (1992)
259 p.
- [17] H. K. Dreiner, H. E. Haber, S. P. Martin, “Two-component spinor techniques and
Feynman rules for quantum field theory and supersymmetry,” Phys. Rept. **494**, 1-196
(2010). [arXiv:0812.1594 [hep-ph]].
- [18] D. Tong, “Lectures on Quantum Field Theory”.
- [19] W. Greiner and J. Reinhardt, Berlin, Germany: Springer (1996) 440 p
- [20] J. Polchinski, “String theory. Vol. 1: An introduction to the bosonic string,” Cambridge,
UK: Univ. Pr. (1998) 402 p
- [21] S. Pokorski, “Gauge Field Theories,” Cambridge, Uk: Univ. Pr. (1987) 394 P. (Cambridge Monographs On Mathematical Physics)