PHYS 215A: Lectures on Quantum Field Theory

Benjamín Grinstein
UCSD

December 10, 2014

Chapter 1

Introduction

1.1 Why QFT?

As a new student of the subject you may ask Why study Quantum Field Theory (QFT)? Or why do we *need* QFT? There are many reasons, and I will explain some.

Pair Creation. For starters, particle quantum mechanics (QM) does not account for pair creation. In collisions of electrons with sufficiently large energies it is occasionally observed that out of the collision an electron and a positron are created (in addition to the originating electrons): $e^-e^- \to e^-e^-e^+e^-$. Sufficiently large energies means kinetic energies larger than the rest mass of the e^-e^+ pair, namely $2m_ec^2$. Not only does single particle QM not account for this but in this case the colliding electrons must be relativistic. So we need a relativistic QM that accounts for particle creation.

But what if we are atomic physicists, and only care about much smaller energies? If we care about high precision, and atomic physicists do, then the fact that pair creation is possible in principle cannot be ignored. Consider the perturbation theory calculation of corrections to the n-th energy level, E_n, of some atomic state in QM:

$$\delta E_n = \langle n|H'|n\rangle + \sum_{k\neq n} \frac{\langle n|H'|k\rangle\langle k|H'|n\rangle}{E_n - E_k} + \cdots$$

The second order term involves *all* states, regardless of their energy. Hence states involving e^+e^- states give corrections of order

$$\frac{\delta E}{E} \sim \frac{\text{atomic energy spacing}}{m_ec^2}.$$

This is, *a priori*, as important as the relativistic corrections from kinematics, $H = \sqrt{(m_ec^2)^2 + (pc)^2} = m_ec^2 + \frac{1}{2}p^2/m_e - \frac{1}{8}p^4/m_e^3c^2 + \cdots$. The first relativistic correction is $H' = -\frac{1}{4}(p/m_ec)^2(p^2/2m_e)$ so the fractional correction $\delta E/E \sim (p/m_ec)^2 \sim (p^2/2m_e)/(m_ec^2)$, just as in pair creation. We are lucky that for the Hydrogen atom the pair creation correction happens to be small. But in general we have no right to neglect it.

Instability of relativistic QM. Let's insist in single particle QM and explore some consequences. Since $H = \sqrt{(m_ec^2)^2 + (pc)^2}$ does not give us a proper differential Schrödinger equation, let's try a wave equation using the square of the energy, $E^2 = (mc^2)^2 + (pc)^2$, and replacing, as usual, $E \to i\hbar\partial/\partial t$ and $\vec{p} \to -i\hbar\nabla$. This leads to the Klein-Gordon wave equation,

$$\left[\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2 + \frac{\mu^2}{c^2}\right]\phi(\vec{x},t) = 0$$

where $\mu = mc^2/\hbar$. Since this is a free particle we look for plane wave solutions,

$$f^+(\vec{x},t) = a_k^+ e^{-i\omega_k t + i\vec{k}\cdot\vec{x}} \qquad \text{and} \qquad f^-(\vec{x},t) = a_k^- e^{i\omega_k t - i\vec{k}\cdot\vec{x}},$$

where $a^\pm$ are constants (independent of $\vec{x}$ and t) and

$$\omega_k = \sqrt{c^2\vec{k}^2 + \mu^2}.$$

The interpretation of f^+ is clear: using $E \to i\hbar\partial/\partial t$ and $\vec{p} \to -i\hbar\nabla$ we see that it has energy $E = \hbar\omega_k$ and momentum $\vec{p} = \hbar k$ and these are related by $E = \sqrt{(m_ec^2)^2 + (pc)^2}$. However, we now also have solutions, f^- with energy $E = -\hbar\omega_k$ (and momentum $\vec{p} = -\hbar k$). These have negative energy. For the free particle this is not a problem since we can start with a particle of some energy, positive or negative, and it will stay at that energy. But as soon as we introduce interactions, say by making the particle charged and giving it minimal coupling to the electromagnetic field, the particle can radiate into a lower energy state, but there is no minimum energy. There is a catastrophic instability.

Dirac proposed an ingenious solution to this catastrophe. Suppose the particle is a fermion, say, an electron. Then start the system from a state in which all the negative energy states are occupied. We call this the "Dirac sea." Since no two fermions can occupy a state with the same quantum numbers, we cannot have an electron with positive energy radiate to become a negative energy state (that state is occupied). However, an energetic photon can interact with an electron with $E < -m_ec^2$ and bump it to a state with $E > m_ec^2$. What results is a state with a positive energy electron and a hole in the "Dirac sea." The hole signifies the absence of negative energy and negative charge, so it behaves as a particle with positive energy and positive charge. That is, as the anti-particle of the electron, the positron. So we conclude we cannot have a single particle QM, we are forced to consider pair creation.

A note in passing: as we will discuss later, relativistic fermions are described by the Dirac equation. But a solution of the Dirac equation necessarily solves the Klein-Gordon equation, so the discussion above is in fact appropriate to the fermionic case.

Klein's Paradox. There is another way to see the need to include anti-particles in the solution to the Klein-Gordon equation. As we stated, only if we include interactions will we see a problem. But our description of interactions with an electromagnetic field was heuristic. We can make it more precise by introducing instead a potential. The simplest case, considered by Klein, is that of a particle in one spatial dimension with a step potential, $V(x) = V_0\,\theta(x)$. Here $\theta(x)$ is the Heavyside step-function

$$\theta(x) = \begin{cases} 1 & x > 0 \\ 0 & x < 0 \end{cases}$$

and $V_0 > 0$ is a constant with units of energy. We look for solutions with a plane wave incident from the left ($x < 0$ with $p = \hbar k > 0$):

$$\begin{aligned} \psi_L(x,t) &= e^{-i\omega t + ik_L x} + Re^{-i\omega t - ik_L x} \\ \psi_R(x,t) &= Te^{-i\omega t + ik_R x} \end{aligned}$$

These solve the Klein-Gordon equation with the shifted energy provided

$$ck_L = \sqrt{\omega^2 - \mu^2} \qquad \text{and} \qquad ck_R = \sqrt{(\omega - \omega_0)^2 - \mu^2}$$

where $\omega_0 = V_0/\hbar$. We now determine the transmission (T) and reflection (T) coefficients matching the solutions at $x = 0$; using $\psi_L(t,0) = \psi_R(t,0)$ and $\frac{\partial \psi_L(t,0)}{\partial t} = \frac{\partial \psi_R(t,0)}{\partial t}$ we get

$$T = \frac{2k_L}{k_L + k_R} \qquad \text{and} \qquad R = \frac{k_L - k_R}{k_L + k_R}.$$

Formally the solution looks just like in the non-relativistic (NR) case. Indeed, for $\omega > \mu + \omega_0$ both wave-vectors k_L and k_R are real, and so are both transmission coefficients and we have a transmitted and a reflected wave. Similarly if $\omega_0 - \mu < \omega < \omega_0 + \mu$ then k_R has a nonzero imaginary part and we have total reflection (the would be transmitted wave is exponentially damped).

But for $V_0 > 2mc^2$, which corresponds to an energy large enough that pair creation is energetically allowed, there is an unusual solution. If $\mu < \omega < \omega_0 - \mu$, which means $\omega - \mu > 0$ and $\omega - \mu < \omega_0 - 2\mu$, we obtain both k_L and k_R are real, and so are both T and R. There is a non-zero probability that the particle, which has less energy than the height of the barrier ($E - mc^2 - V_0 < 0$), is transmitted. This weird situation, called "Klein's paradox," can be understood in terms of pair creation. A complete treatment of the problem requires (as far as I know) a fully quantum field theoretic treatment, so for that we will have to wait until later in this course. But the result can be easily described: the incident particle is fully reflected, but is accompanied by particle-antiparticle pairs.

Bohr's Box. Above we talked about "localizing" a particle. Let's make this a bit more precise (not much, though). Suppose that in order to localize a free particle we put it in a large box. We don;t know where the particle is other than that it is inside the box. If choose the box large enough, of sides $L_{x,y,z} \gg \lambda_C$, then the uncertainly in its momentum can be small, $\Delta p_{x,y,z} \gtrsim \hbar/L_{x,y,z}$. Then the kinetic energy can be small, $E = \frac{\vec{p}^2}{2m} \sim \frac{\hbar^2}{2m}(L_x^{-2} + L_y^{-2} + L_z^{-2})$. Now suppose one side of the box, parallel to the yz plane, is movable (the box is a cylinder, the movable box is a piston). So we can attempt to localize the particle along the x axis by compressing it, that is, by decreasing L_x. Once $L_x \sim \lambda_C$ the uncertainty in the energy of the particle is $\Delta E \sim mc^2$. Now o localize the particle we have introduced interaction, those of the particle with the walls of the box that keeps it contained. But as we have seen when the energy available exceeds $2mc^2$ interactions require a non-vanishing probability for pair creation. We conclude that as we try to localize the particle to within a Compton wavelength, $\lambda_C = \hbar/mc$, we get instead a state which is a combination of the particle plus a particle-antiparticle pair. Or perhaps two pairs, or three pairs, or .... In trying to localize a particle not only we have lost certainty on its energy, as is always the case in QM, but we have also lost certainty on the number of particles we are trying to localize!

This gives us some physical insight into Klein's paradox. The potential step localizes the particle over distances smaller than a Compton wavelength. The uncertainty principle then requires that the energy be uncertain by more than the energy required for pair production. It can be shown that if the step potential is replaced by a smooth potential that varies slowly between 0 and V_0 over distances d larger than λ_C then transmission is exponentially damped, but as d is made smaller than λ_C Klein's paradox re-emerges (Sauter).

Democracy. It is well known that elementary particles are indistinguishable: every electron in the world has the same mass, charge and magnetic moment. In NRQM we account for this in an ad-hoc fashion. We write

$$H = \sum_i \frac{\vec{p}_i^2}{2m_e} + \frac{e}{m_e c}\vec{A}(\vec{x}_i) \cdot \vec{p}_i + \vec{B}(\vec{x}_i) \cdot \vec{\mu}_i \qquad \text{where } \vec{\mu}_i = \frac{ge\hbar}{2m_e c}\vec{\sigma}_i,$$

but we have put in by hand that the mass m_e, the charge e and the gyromagnetic ratio g are the same for all electrons. Where does this democratic choice come from?

Moreover, photons, which are quanta of the electromagnetic field, are introduced by second quantizing the field. But electrons are treated differently. Undemocratic! If instead we insist that all particles are quanta of field excitations not only we will have a more democratic (aesthetically pleasing?) setup but we will have an explanation for indistinguishability, since all corpuscular excitations of a field carry the same quantum numbers automatically.

Indistinguishability is fundamentally importance in Pauli's exclusion principle. If several electrons had slightly different masses or slightly different charges from each other, then they could be distinguished and they could occupy the same atomic orbital (which in fact would not be the same, but slightly different). More generally, indistinguishability is at the heart of "statistics" in QM. But the choice of Bose-Einstein vs Fermi-Dirac statistics is a recipe in particle QM, it has to be put in by hand. Since QFT will give us indistinguishability automatically, one may wonder if it also has something to say about Bose-Einstein vs Fermi -Dirac statistics. In fact it does. We will see that consistent quantization of a field describing spin-0 corpuscles requires them to obey Bose-Einstein statistics, while quantization of fields that give spin-1/2 particles results in Fermi-Dirac statistics.

Causality. In relativistic kinematics faster than light signal propagation leads to paradoxes. The paradoxes come about because faster than light travel violates our normal notion of causality. A spaceship (it's always a spaceship) moving faster than light from event A to event B is observed in other frames as moving from B to A. You never had to worry about this in NRQM because you never had to worry about it in NR classical mechanics. But you do worry about it in relativistic mechanics so you must be concerned that related problems arise in relativistic QM.

In particle QM we can define an operator $\hat{\vec{x}}$ and we can use eigenstates of this operator, $\hat{\vec{x}}|\vec{x}\rangle = \vec{x}|\vec{x}\rangle$ to describe the particle. For now I will use a hat to denote operators on the Hilbert space, but I will soon drop this. The operator is conjugate to the momentum operator, $\hat{\vec{p}}$, in the sense that $i[\hat{x}_j, \hat{p}_k] = \delta_{jk}$, and this gives rise to a relation between their eigenstates,

$$\langle\vec{p}|\vec{x}\rangle = \frac{1}{(2\pi\hbar)^{3/2}} e^{-i\vec{p}\cdot\vec{x}/\hbar}.$$

A state $|\psi\rangle$ at $t = 0$ evolves into $e^{-i\hat{H}t/\hbar}|\psi\rangle$ a time t later. We can ask what does the state of a particle localized at the origin at $t = 0$ evolve into a time t later. Now, in NRQM we know the answer: the particle spreads out. Whether it spreads out faster than light or not is not an issue so we don't ask the question. But in relativistic QM it matters, so we address this. Note however that for consistency we must use a relativistic Hamiltonian. For a free particle we can use $\hat{H} = \sqrt{(\hat{\vec{p}}c)^2 + (mc^2)^2}$ since its action on states is unambiguously given by its action on momentum eigenstates, $\hat{H}|\vec{p}\rangle = \sqrt{(\vec{p}c)^2 + (mc^2)^2}|\vec{p}\rangle$. The probability amplitude of finding the particle at $\vec{x}$ at time t (given that it started as a localized state at the origin at time $t = 0$) is

$$\begin{aligned}
\langle\vec{x}|e^{-i\hat{H}t/\hbar}|\vec{x}=0\rangle &= \int d^3p\,\langle\vec{x}|\vec{p}\rangle\langle\vec{p}|e^{-i\hat{H}t/\hbar}|\vec{x}=0\rangle \qquad \text{(complete set of states)} \\
&= \int d^3p\,\frac{1}{(2\pi\hbar)^3}e^{i\vec{p}\cdot\vec{x}/\hbar}e^{-iE_pt/\hbar} \qquad \text{(where } E_p = \sqrt{(pc)^2+(mc^2)^2}\text{)} \\
&= \int_0^\infty \frac{p^2dp}{(2\pi\hbar)^3}\int_{-1}^{1} d\cos\theta \int_0^{2\pi} d\phi\, e^{ipr\cos\theta/\hbar}e^{-iE_pt/\hbar} \qquad (p=|\vec{p}|;\ r=|\vec{x}|) \\
&= -\frac{i}{(2\pi\hbar)^2 r}\int_{-\infty}^{\infty} p\,dp\,e^{irp/\hbar}e^{-iE_pt/\hbar} \qquad (1.1)
\end{aligned}$$

This is not an easy integral to compute. But we can easily show that it does not vanish for $r > ct > 0$. This means that there is a non-vanishing probability of finding the particle at places that require it propagated faster than light to get there in the allotted time. This is a violation of causality.

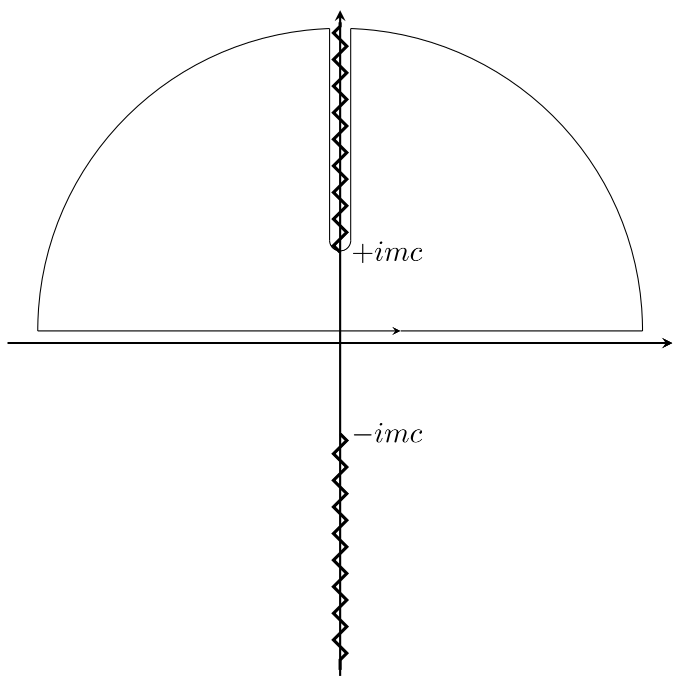


Figure 1.1: Contour integral for evaluating (1.1).

Before discussing this any further let's establish the claim that the integral does not vanish for $r/t > c$. The integral is difficult to evaluate because it is oscillatory. However, we can use complex analysis to relate the integral to one performed over purely imaginary momentum, turning the oscillating factor into an exponential convergence factor. So consider the analytic structure of the integrand. Only the square root defining the energy is non-analytic, with a couple of branch points at $p = \pm imc$. Choose the branch cut to extend from $+imc$ to $+i\infty$ and from $-imc$ to $-i\infty$ along the imaginary axis; see Fig. 1.1. The integral over the contour C vanishes (there are no poles of the integrand), and for $r > ct$ the integral over the semicircle of radius R vanishes exponentially fast as $R \to \infty$. Then the integral we want is related to the one on both sides of the positive imaginary axis branch cut,

$$\langle \vec{x}| e^{-i\hat{H}t/\hbar} |\vec{x} = 0\rangle = \frac{i}{(2\pi\hbar)^2 r} \int_{mc}^{\infty} p\, dp\, e^{-rp/\hbar} \left(e^{\sqrt{p^2-(mc)^2}ct/\hbar} - e^{-\sqrt{p^2-(mc)^2}ct/\hbar} \right). \tag{1.2}$$

The integrand is everywhere positive. It decreases exponentially fast as r increases for fixed t, so the violation to causality is small. But any violation to causality is problematic.

1.2 Units and Conventions

You surely noticed the proliferation of c and $\hbar$ in the equations above. The play no role, other than to keep units consistent throughout. So for the remainder of the course we will adopt units in which $c = 1$ and $\hbar = 1$. You are probably familiar with $c = 1$ already: you can measure distance in light-seconds and then x/t has no units. But now, in addition energy momentum and mass and measured in the same units (after all $E^2 = (pc)^2 + (mc^2)^2$). We denote units by square brackets, the units of X are $[X]$; we have $[E] = [p] = [m]$, and $[x] = [t]$.

The choice $\hbar = 1$ may be less familiar. Form the uncertainty condition, $\Delta p \Delta x \geq \hbar$, so $[p] = [x^{-1}]$. This together with the above (from $c = 1$) means that everything can be measured in units of energy, $[E] = [p] = [m] = [x^{-1}] = [t^{-1}]$. In particle physics it is customary to use GeV as the common unit. That's because many elementary particles have masses of the order or a GeV:

particle	symbol	mass(GeV)
proton	m_p	0.938
neutron	m_n	0.940
electron	m_e	5.11×10^{-4}
W-boson	M_W	80.4
Z-boson	M_Z	91.2
higgs boson	M_h	126

To convert units, use $\hbar = 6.582 \times 10^{-25}$ GeV·sec, and often conveniently $\hbar c = 0.1973$ GeV·fm, where fm is a Fermi, or femtometer, 1 fm $= 10^{-15}$ m, a typical distance scale in nuclear physics.

In these units the fine structure constant, $\alpha = e^2/4\pi\hbar c$ is a pure number,

$$\alpha = \frac{e^2}{4\pi} \approx \frac{1}{137}.$$

Since we will study relativistic systems it is useful toset up conventions for our notation. We use the "mostly-minus" metric, $\eta_{\mu\nu} = \text{diag}(+,-,-,-)$. That is, the invariant interval is $ds^2 = \eta_{\mu\nu}dx^\mu dx^\nu$. The Einstein convention, an implicit sum over repeated indices unless otherwise stated, is adopted. Four-vectors have upper indices, $a = (a^0, a^1, a^2, a^3)$, and the dot product is

$$a \cdot b = \eta_{\mu\nu}a^\mu b^\nu = a^0b^0 - a^1b^1 - a^2b^2 - a^3b^3 = a^0b^0 - \vec{a} \cdot \vec{b} = a^0b^0 - a^ib^i$$

We use latin indices for the spatial component of the 4-vectors, and use again the Einstein convention for repeated latin indices, $\vec{a} \cdot \vec{b} = a^ib^i$. Indices that run from 0 to 3 are denoted by greek letters. We also use $a^2 = a \cdot a$ and $|\vec{a}\,|^2 = \vec{a}^{\,2} = \vec{a} \cdot \vec{a}$. Sometimes we even use $a^2 = \vec{a} \cdot \vec{a}$, even when there is a 4-vector a^μ; this is confusing, and should be avoided, but when it is used it is always clear from the context whether a^2 is the square of the 4-vector or the 3-vector.

The inverse metric is denoted by $\eta^{\mu\nu}$,

$$\eta^{\mu\nu}\eta_{\nu\lambda} = \delta^\mu_\lambda$$

where δ^μ_ν is a Kronecker-delta, equal to unity when the indices are equal, otherwise zero. Numerically in Cartesian coordinates $\eta^{\mu\nu}$ is the same matrix as the metric $\eta_{\mu\nu}$, $\eta^{\mu\nu} = \text{diag}(+,-,-,-)$, but it is convenient to differentiate among them because they need not be the same in other coordinate systems. We can use the metric and inverse metric to define lower index vectors (I will not use the names "covariant" and "contravariant") and to convert among them:

$$a_\mu = \eta_{\mu\nu}a^\nu, \qquad a^\mu = \eta^{\mu\nu}a_\nu.$$

Then the dot product can be expressed as

$$a \cdot b = \eta_{\mu\nu}a^\mu b^\nu = a^\mu b_\mu = a_\mu b^\mu.$$

Generalized Einstein convention: any type of repeated index is understood as summed, unless explicitly stated. For example, if we have a set of quantities ϕ^a with $a = 1, \ldots, N$, then $\phi^a\phi^a$ stands for $\sum_{a=1}^N \phi^a\phi^a$.

1.3 Lorentz Transformations

Lorentz transformations map vectors into vectors

$$a^\mu \to \Lambda^\mu{}_\nu a^\nu$$

preserving the dot product,

$$a \cdot b \to (\Lambda a) \cdot (\Lambda b) = a \cdot b$$

Since this must hold for any vectors a and b, we must have

$$\Lambda^\lambda{}_\mu \Lambda^\sigma{}_\nu \eta_{\lambda\sigma} = \eta_{\mu\nu} \tag{1.3}$$

Multiplying by the inverse metric, $\eta^{\nu\rho}$

$$\Lambda^\lambda{}_\mu \Lambda^\sigma{}_\nu \eta_{\lambda\sigma} \eta^{\nu\rho} = \delta^\rho_\mu$$

we see that

$$\Lambda_\lambda{}^\rho \equiv \Lambda^\sigma{}_\nu \eta_{\lambda\sigma} \eta^{\nu\rho}$$

is the inverse of $\Lambda^\lambda{}_\rho$, $(\Lambda^{-1})^\rho{}_\lambda = \Lambda_\lambda{}^\rho$. Eq. (1.3) can be written in matrix notation as

$$\Lambda^T \eta \Lambda = \eta \tag{1.4}$$

where the superscript "T" stands for "transpose,"

$$(\Lambda^T)_\mu{}^\lambda = \Lambda^\lambda{}_\mu.$$

Lorentz transformations form a group with multiplication given by composition of transformations (which is just matrix multiplication, $\Lambda_1 \Lambda_2$): there is an identity transformation (the unit matrix), an inverse to every transformation (introduced above) and the product of any two transformations is again a transformation. The *Lorentz group* is denoted $O(1,3)$. Taking the determinant of (1.4), and using $\det(AB) = \det(A)\det(B)$ and $\det(A^T) = \det(A)$ we have $\det^2(\Lambda) = 1$, and since Λ is real, $\det(\Lambda) = +1$ or -1. The product of two Lorentz transformations with $\det(\Lambda) = +1$ is again a Lorentz transformation with $\det(\Lambda) = +1$, so the set of transformations with $\det(\Lambda) = +1$ form a subgroup, the group of *Special (or Proper) Lorentz Transformations*, $SO(1,3)$. Among the $\det(\Lambda) = -1$ transformations is the parity transformation, that is, reflection aboutt he origin, $\Lambda = \text{diag}(+1, -1, -1, -1)$.

Taking $\mu = \nu = 0$ in (1.3) we have

$$(\Lambda^0{}_0)^2 = 1 + \sum_{i=1}^{3} (\Lambda^i{}_0)^2 > 1$$

So any Lorentz transformation has either ${\Lambda^0}_0 \geq 1$ or ${\Lambda^0}_0 \leq 1$. The set of transformations with ${\Lambda^0}_0 \geq 1$ are continuously connected, and so are the ones with ${\Lambda^0}_0 \leq 1$, but no continuous transformation can take one type to the other. Among those with ${\Lambda^0}_0 \geq 1$ is the identity transformation; among those with ${\Lambda^0}_0 \leq 1$ is time reversal, $\Lambda = \mathrm{diag}(-1,+1,+1,+1)$. We will have much more to say about parity and time reversal later in this course. Transformations with ${\Lambda^0}_0 \geq 1$ are called *orthochronous* and they form a subgroup denoted by $O^+(1,3)$, The subgroup of proper, orthochronous transformations, sometimes called the *restricted Lorentz group*, $SO^+(1,3)$.

Examples of Lorentz transformations: boosts along the x-axis

$$\Lambda = \begin{pmatrix} \cosh\theta & \sinh\theta & 0 & 0 \\ \sinh\theta & \cosh\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.5}$$

and rotations

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix}$$

where R is a 3×3 matrix satisfying $R^T R = 1$. Rotations form a group, the group of 3×3 orthogonal matrices, $O(3)$. Note that $\det^2(R) = 1$, so the matrices with $\det(R) = +1$ form a subgroup, the group of Special Orthogonal transformations, $SO(3)$.

1.3.1 More conventions

We will use a common shorthand notation for derivatives:

$$\partial_\mu \phi = \frac{\partial\phi}{\partial x^\mu}$$

Note that the lower index on ∂_μ goes with the upper index in the "denominator" in $\frac{\partial\phi}{\partial x^\mu}$. We can see that this works correctly by taking

$$\frac{\partial}{\partial x^\mu}(x^\nu a_\nu) = a_\mu,$$

an lower index vector. Of course, the justification is how the derivative transforms under a Lorentz transformation. If $(x')^\mu = {\Lambda^\mu}_\nu x^\nu$ then

$$\frac{\partial}{\partial x'^\mu} = \frac{\partial x^\lambda}{\partial x'^\mu}\frac{\partial}{\partial (x)^\lambda} = \frac{\partial}{\partial x'^\mu}\left({(\Lambda^{-1})^\lambda}_\nu (x')^\nu\right)\frac{\partial}{\partial (x)^\lambda} = {(\Lambda^{-1})^\lambda}_\mu \frac{\partial}{\partial (x)^\lambda}$$

as it should: $\partial'_\mu = (\Lambda^{-1})^\lambda{}_\mu \partial_\lambda$. Sometimes we use also $\partial^\mu = \eta^{\mu\nu}\partial_\nu$. For integrals we use standard notation,

$$\int dt\, dx\, dy\, dz = \int dx^0\, dx^1\, dx^2\, dx^3 = \int d^4x = \int dt \int d^3x.$$

This is a Lorentz invariant (because the Jacobian, $|\det(\Lambda)| = 1$):

$$\int d^4x' = \int d^4x.$$

1.4 Relativistic Invariance

What does it mean to have a relativistic formulation of QM? A QM system is completely defined by its states (the Hilbert space $\mathcal{H}$) and the action of the Hamiltonian $\hat{H}$ on them. Consider again a free particle. Since $H = E$ is in a 4-vector with $\vec{p}$ we define a QM system by

$$\hat{\vec{p}}|\vec{p}\rangle = \vec{p}|\vec{p}\rangle, \qquad \hat{H}|\vec{p}\rangle = \sqrt{\vec{p}^2 + m^2}|\vec{p}\rangle.$$

What do we mean by this being relativistic? That is, how is this invariant under Lorentz transformations?

To answer this it is convenient to first review rotational invariance, which we are more familiar with. In QM for each rotation R there is an operator on $\mathcal{H}$, $\hat{U}(R)$ such that

$$\hat{U}(R)|\vec{p}\rangle = |R\vec{p}\rangle$$

I will stop putting "hats" on operators when it is clear we are speaking of operators. So from here on $U(R)$ stands for $\hat{U}(R)$, etc. Note that

$$\begin{aligned}
U(R)\hat{\vec{p}}\,|\vec{p}\,\rangle = \vec{p}\,U(R)|\vec{p}\rangle = \vec{p}\,|R\vec{p}\rangle \\
\Rightarrow U(R)\hat{\vec{p}}\,U(R)^{-1}|R\vec{p}\,\rangle = R^{-1}(R\vec{p}\,|R\vec{p}\,\rangle) = R^{-1}\hat{\vec{p}}\,|R\vec{p}\,\rangle \\
\Rightarrow U(R)\hat{\vec{p}}\,U(R)^{-1} = R^{-1}\hat{\vec{p}} \quad (1.6)
\end{aligned}$$

We would like U to be unitary, so that probability of finding a state is the same as that of finding the rotated state (that, and $[U, H] = 0$, is what we mean by a symmetry). We should be able to prove that $U^\dagger U = UU^\dagger = 1$. Let's assume the states $|R\vec{p}\,\rangle$ are normalized by

$$\langle\vec{p}\,'|\vec{p}\rangle = \delta^{(3)}(\vec{p}\,' - \vec{p}).$$

Then to show $UU^\dagger = 1$ we use a neat trick:

$$U(R)U^\dagger(R) = U(R)\int d^3p\,|\vec{p}\rangle\langle\vec{p}|U(R)^\dagger = \int d^3p\,|R\vec{p}\rangle\langle R\vec{p}| = \int d^3p'\,|\vec{p}\,'\rangle\langle\vec{p}\,'| = 1,$$

where we have used $\vec{p}\,' = R\vec{p}$ and $d^3p = d^3p'$. The latter is non-trivial. It reflects the (clever) choice of normalization of states, which leads to a rotational invariant measure.

You can see things go wrong if one defines the normalization of states differently. Say $|\vec{p}\,\rangle_B = (1+|\vec{a}\cdot\vec{p}\,|)|\vec{p}\,\rangle$ where $\vec{a}$ is a fixed vector (and the subscript "B" stands for "Bad"). Then

$$ {}_B\langle\vec{p}\,'|\vec{p}\rangle_B = (1+|\vec{a}\cdot\vec{p}\,|)^2\delta^{(3)}(\vec{p}\,'-\vec{p}) \Rightarrow 1 = \int d^3p \frac{|R\vec{p}\rangle_{B\,B}\langle R\vec{p}|}{(1+|\vec{a}\cdot\vec{p}\,|)^2}. $$

The point is that there is a choice of states for which we can show easily $UU^\dagger = 1$. It is true that $UU^\dagger = 1$ even for the bad choice of states, it is just more difficult to prove.

We still have to show that U commutes with H, but this is simple:

$$ \begin{aligned} U(R)H|\vec{p}\,\rangle = \sqrt{\vec{p}^{\,2}+m^2}|\vec{p}\,\rangle \Rightarrow U(R)HU(R)^{-1}|R\vec{p}\,\rangle &= \sqrt{(R\vec{p})^2+m^2}|R\vec{p}\,\rangle \\ &\Rightarrow U(R)HU(R)^{-1} = H. \end{aligned} $$

where we used $(R\vec{p})^2 = \vec{p}^{\,2}$ in the second step.

Now consider Lorentz Invariance: $p \to \Lambda p$ with $\Lambda \in O(1,3)$. Actually, for now we will restrict attention to transformations $\Lambda \in SO(1,3)$ with ${\Lambda^0}_0 \geq 1$. That's because time reversal and spatial inversion (parity) present their own subtleties about which we will have much to say later in the course. As before we have $U(\Lambda)|E,\vec{p}\,\rangle = |\Lambda(E,\vec{p}\,)\rangle$. Note, first, that there is no need to specify E in addition to $\vec{p}$; we are being explicit to understand how Λ acts on states. In fact, since $E^2 - \vec{p}^{\,2} = m^2$, E and $\vec{p}$ fall on a hyperboloid. The action of Λ on $(E,\vec{p})$ just moves points around in the hyperboloid. In order to show $UU^\dagger = 1$ we try the same trick:

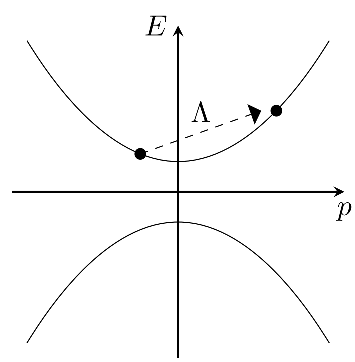


Figure 1.2: The hyperboloid $E^2 - \vec{p}^{\,2} = m^2$; upper(lower) branch has $E > 0 (< 0)$.

$$ U(\Lambda)U^\dagger(\Lambda) \;=\; U(\Lambda)\int d^3p\,|E,\vec{p}\rangle\langle E,\vec{p}|\;\; U(\Lambda)^\dagger \;=\; \int d^3p\,|\Lambda(E,\vec{p})\rangle\langle\Lambda(E,\vec{p})|. $$

But now we hit a snag: if $(E',\vec{p}\,') = \Lambda(E,\vec{p})$, then $d^3p' \neq d^3p$. This is easily seen by considering boosts along the x direction, Eq. (1.5). But our experience from the discussion above suggests we find a better basis of states, one chosen so that the measure of integration is Lorentz invariant. In fact, we can engineer back the normalization of states from requiring a invariant measure. Start from the observation that the 4-dimensional measure is invariant, $d^4p' = d^4p$. Now pick from this the upper hyperboloid in Fig. 1.2 in a manner that explicitly preserves Lorentz invariance. This can be done using a delta function, $\delta(p^2 - m^2) = \delta((p^0)^2 - \vec{p}^2 - m^2)$ and a step function $\theta(p^0)$ to select the upper solution of the δ-function constraint. Note that $\theta(p^0)$ is invariant under transformations with ${\Lambda^0}_0 \geq 1$. So we take our measure to be

$$\begin{aligned} d^4p\ \delta(p^2 - m^2)\theta(p^0) &= d^4p\ \delta((p^0)^2 - E_{\vec{p}}^2)\theta(p^0) \quad (\text{where } E_{\vec{p}} \equiv \sqrt{\vec{p}^2 + m^2}) \\ &= d^3p\ dp^0\ \frac{1}{2E_{\vec{p}}}\delta(p^0 - E_{\vec{p}}) \\ &= \frac{d^3p}{2E_{\vec{p}}} \end{aligned}$$

For later convenience we introduce a constant factor and compact notation,

$$(dp) = \frac{d^3p}{(2\pi)^3 2E_{\vec{p}}}\,.$$

The corresponding (relativistic) normalization of states is

$$\langle \vec{p}\,' | \vec{p} \rangle = (2\pi)^3 2E_{\vec{p}}\ \delta^{(3)}(\vec{p}\,' - \vec{p}) \tag{1.7}$$

Now we have,

$$\begin{aligned} U(\Lambda)U^\dagger(\Lambda) = U(\Lambda)\int (dp)\, |\vec{p}\rangle\langle\vec{p}|\, U(\Lambda)^\dagger &= \int (dp)\, |\Lambda(E,\vec{p}\,)\rangle\langle\Lambda(E,\vec{p})| \\ &= \int (dp)\, |(E',\vec{p}\,')\rangle\langle(E',\vec{p}\,')| = \int (dp')\, |\vec{p}\,'\rangle\langle\vec{p}\,'| = 1. \end{aligned}$$

From Eq. (1.7) one can easily show $U^\dagger U = 1$. Also,

$$U(\Lambda)(H,\hat{\vec{p}}\,)U(\Lambda)^{-1} = \Lambda^{-1}(H,\hat{\vec{p}}\,)$$

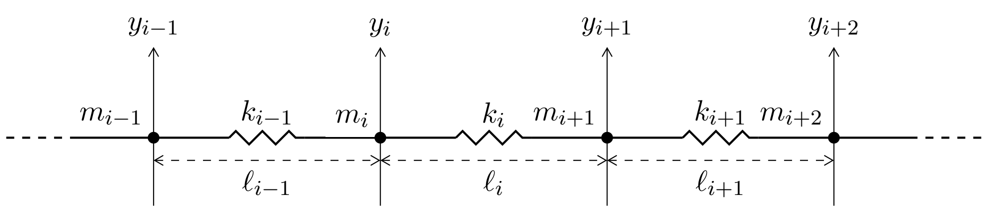
which is what we mean by relativistic co-variance.

Chapter 2

Field Quantization

2.1 Classical Fields

Consider a (classical, non-relativistic) system of masses, m_i, connected by springs, so that m_i and m_{i+1} are connected by a spring with spring constat k_i. In equilibrium the masses all lie on a straight line, and the distance between masses m_i and m_{i+1} is ℓ_i. The masses are free to move only on a fixed direction perpendicular to this straight line. This is shown in the figure below. We are free to use



a coordinate system to describe the positions of the masses with the x-axis along the equilibrium straight line and the y-axis the transverse direction in which the masses are constrained to move. The i-th mass has coordinates $\vec{x}_i$. To describe the dynamics of this system we construct the Lagrangian,

$$L = L(\dot{y}_i, y_i) = \sum_i \tfrac{1}{2} m_i \dot{y}_i^2 - V(|\vec{x}_{i+1} - \vec{x}_i|)$$

where

$$V = \sum_i \tfrac{1}{2} k_i |\vec{x}_{i+1} - \vec{x}_i|^2 = \tfrac{1}{2} k_i \left((y_{i+1} - y_i)^2 + \ell_i^2 \right)$$

so that, dropping the irrelevant constant

$$L = L(\dot{y}_i, y_i) = \sum_i \tfrac{1}{2} m_i \dot{y}_i^2 - \tfrac{1}{2} k_i (y_{i+1} - y_i)^2 \,.$$

We are interested in this system in the limit that we cannot resolve the individual masses, so by our measuring apparatus the system appears as a continuum. Mathematically we want to take the limit $\ell_i \to 0$ and describe the displacement of the system from the x-axis at some point x along the axis, at time t, by a function $\xi(x,t)$. This function is called a *field*. Since the displacement is y_i at position x_i along the axis, we identify $y_i(t) \to \xi(x_i,t)$. Note that while x_i is used classically as a coordinate of a particle, in $\xi(x,t)$ it is just a label telling us where we are measuring the displacement ξ. This is an important point, so I dwell on it a bit, since first time students of quantum field theory often get confused with the role of x in the argument of a field. The field value itself can be measuring something other than displacement. For example, it could be temperature or pressure, or electric field. The argument x, or in the three-dimensional case $\vec{x}$ of a field indicates where the field has a particular value. So x (or $\vec{x}$) is not a dynamical variable, but ξ is.

To rewrite the Lagrangian in terms of the field, use

$$y_{i+1} - y_i \to \xi(x_i + \ell_i, t) - \xi(x_i, t) = \ell_i \left.\frac{\partial \xi}{\partial x}\right|_{x=x_i} + \cdots ,$$

where the ellipses stand for terms with higher powers of ℓ_i, and $\dot{y}_i \to \left.\frac{\partial \xi}{\partial t}\right|_{x=x_i}$. Multiplying and dividing by ℓ_i and interpreting $\ell_i = x_{i+1} - x_i$ as the Δx, we have then

$$L = \sum_i \Delta x \left[\frac{1}{2} \frac{m_i}{\ell_i} \left(\frac{\partial \xi}{\partial t}\right)^2 - \frac{1}{2} k_i \ell_i \left(\frac{\partial \xi}{\partial x}\right)^2 \right]$$

where the derivatives are evaluated at $x = x_i$. We take the limit $\ell_i \to 0$ keeping the ratio m_i/ℓ_i and the product $k_i \ell_i$ finite. These fixed ratio and product then can be characterized with functions $\sigma(x)$ and $\kappa(x)$ (with $\sigma(x_i) = m_i/\ell_i$ and $\kappa(x_i) = k_i \ell_i$ in the limiting process. Of course, the sum becomes an integral and we have

$$L = \int dx \, \mathcal{L} = \int dx \, \left[\frac{1}{2} \sigma(x) \left(\frac{\partial \xi}{\partial t}\right)^2 - \frac{1}{2} \kappa(x) \left(\frac{\partial \xi}{\partial x}\right)^2 \right] .$$

The function $\mathcal{L} = \mathcal{L}(\partial_t \xi, \partial_x \xi, \xi, x, t)$ is called a *Lagrangian density*. This is our first example of a field theory. The dynamics of the field $\xi(x,t)$ is specified by the Lagrangian density

$$\mathcal{L} = \frac{1}{2} \sigma(x) \left(\frac{\partial \xi}{\partial t}\right)^2 - \frac{1}{2} \kappa(x) \left(\frac{\partial \xi}{\partial x}\right)^2 .$$

In no time we will get tired of saying "Lagrangian density" so, as is commonly done in practice, we will improperly refer to $\mathcal{L}$ as a Lagrangian. The distinction should be clear from the context (if it is integrated it is actually a Lagrangian, else it is a density). It should be no surprise that a dynamical variable that varies continuously in space requires densities for its description.

We are often interested in systems that are homogeneous in space, that is, the location of the origin of the coordinate system should be irrelevant. So we impose that the Lagrangian be invariant under a space translation $x' = x - a$. The fields change into new fields $\xi'(x', t) = \xi(x, t)$, which is just a relabeling of the dynamical variables (a canonical transformation, to be precise). But $\sigma(x)$ and $\kappa(x)$ do change, unless $\sigma(x + a) = \sigma(x)$ and $\kappa(x + a) = \kappa(x)$ for any a. This implies, $\sigma(x) = \sigma =$ constant and $\kappa(x) = \kappa =$ constant. Given this, it is convenient to introduce a change of variables, $\phi(x, t) = \sqrt{\kappa}\xi(x, t)$, so that the Lagrangian density is written more simply:

$$\mathcal{L} = \frac{1}{2}\frac{1}{c_s^2}\left(\frac{\partial\phi}{\partial t}\right)^2 - \frac{1}{2}\left(\frac{\partial\phi}{\partial x}\right)^2 . \tag{2.1}$$

where we have introduced the shorthand $c_s^2 = \kappa/\sigma$.

Before we go over to find the equations of motion for this system, let's review the derivation of the Euler-Lagrange equations (or *equations of motion*) for a system with discrete degrees of freedom. Given a Lagrangian $L = L(\dot{q}^a, q^a)$, Hamilton's principle says the equations of motion are obtained from requiring that the action integral be extremized:

$$\delta S[q^a(t)] = \delta \int_{t_1}^{t_2} dl\, L(\dot{q}^a, q^a) = 0 \qquad \text{with} \quad q^a(t_1) = q^a_{\text{ini}} \text{ and } q^a(t_2) = q^a_{\text{fin}}$$

Computing we have,

$$0 = \int_{t_1}^{t_2} dt \sum_a \left[\frac{\partial L}{\partial \dot{q}^a}\frac{d}{dt}\delta q^a + \frac{\partial L}{\partial q^a}\delta q^a\right] = \int_{t_1}^{t_2} dt \sum_a \delta q^a \left[-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^a} + \frac{\partial L}{\partial q^a}\right] + \frac{\partial L}{\partial \dot{q}^a}\delta q^a\Big|_{t_1}^{t_2}$$

The last term vanishes by the fixed boundary conditions $\delta q^a(t_{1,2}) = 0$, and the first term vansihes for arbitrary variation $\delta q^a(t)$ if

$$\frac{\partial L}{\partial q^a} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^a} + = 0$$

These are the Euler-Lagrange equations.

Moving on to the continuum case, we apply the same principle, that the action integral be an extremum under variations of the dynamical variable, $\phi(x, t)$:

$$\delta S[\phi(x, t)] = \delta \int dt L = \delta \int_{t_1}^{t_2} dt \int dx \mathcal{L}(\partial_t\phi, \partial_x\phi, \phi, t) = 0.$$

The boundary conditions are now $\phi(x, t_1) = \phi_{\text{ini}}(x)$ and $\phi(x, t_2) = \phi_{\text{fin}}(x)$. We have intentionally not specified boundary conditions for the x-integration. This will allow us to decide what are reasonable conditions as we derive equations of motion. Computing the variation of S does not introduce new complications:

$$\begin{aligned} 0 = \delta S &= \int_{t_1}^{t_2} dt \int dx \left[\frac{\partial \mathcal{L}}{\partial \left(\frac{\partial \phi}{\partial t}\right)} \frac{\partial \delta\phi}{\partial t} + \frac{\partial \mathcal{L}}{\partial \left(\frac{\partial \phi}{\partial x}\right)} \frac{\partial \delta\phi}{\partial x} + \frac{\partial \mathcal{L}}{\partial \phi}\delta\phi \right] \\ &= \int_{t_1}^{t_2} dt \int dx \delta\phi \left[-\frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \left(\frac{\partial \phi}{\partial t}\right)} - \frac{\partial}{\partial x}\frac{\partial \mathcal{L}}{\partial \left(\frac{\partial \phi}{\partial x}\right)} + \frac{\partial \mathcal{L}}{\partial \phi} \right] \\ &\qquad + \int dx \frac{\partial \mathcal{L}}{\partial \left(\frac{\partial \phi}{\partial t}\right)} \delta\phi \Big|_{t_1}^{t_2} + \int_{t_1}^{t_2} dt \frac{\partial \mathcal{L}}{\partial \left(\frac{\partial \phi}{\partial x}\right)} \delta\phi \Big|_{x=?} \end{aligned}$$

The first term on the last line vanishes by the boundary conditions at $t = t_{1,2}$. The second term vanishes if we fix boundary conditions on $\phi(x,t)$ at the limits of integration for x. If the field is defined over the whole line $x \in (-\infty, \infty)$ then we can specify $\lim_{x\to\pm\infty} \phi(x,t) = 0$. This is reasonable. If you start with the collection of springs and masses form its equilibrium configuration, and poke it somewhere, it will take infinite time for the masses infinitely far away to be excited. But we are considering finite time, and in finite time the masses far away never get displaced from equilibrium. This is even clearer in the continuum case. We will see shortly that ϕ satisfies a wave equation with finite speed of propagation. Alternatively, we can imagine the case of a finite system of springs and masses extending from $x = 0$ to $x = L$. The limit of $\ell_i \to 0$ still requires that we take the number of masses and springs to infinity, but we can do so with the field confined to the region $x \in [0, L]$. In this case we need to introduce boundary conditions at $x = 0$ and L. If the ends of the line of masses are fixed, then in the limit $\phi(0,t)$ $\phi(L,t)$ are fixed. Another popular setup is to have periodic boundary conditions, $\phi(L,t) = \phi(0,t)$. This means the field is defined on a 1-dimensional torus (really a circle, but the generalization to higher dimensions is a torus). This also makes the last terms vanish. Physically, if we have only finite time and the size of the system L is sufficiently large, the precise choice of boundary conditions should be irrelevant.

Setting to zero the coefficient of the arbitrary variation $\delta\phi(x,t)$ gives the Euler-Lagrange equations:

$$-\frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \left(\frac{\partial \phi}{\partial t}\right)} - \frac{\partial}{\partial x}\frac{\partial \mathcal{L}}{\partial \left(\frac{\partial \phi}{\partial x}\right)} + \frac{\partial \mathcal{L}}{\partial \phi} = 0\,,$$

To obtain equations of motion in our example, (2.1), compute,

$$\frac{\partial \mathcal{L}}{\partial \phi} = 0\,, \qquad \frac{\partial \mathcal{L}}{\partial \left(\frac{\partial \phi}{\partial x}\right)} = -\frac{\partial \phi}{\partial x}\,, \qquad \frac{\partial \mathcal{L}}{\partial \left(\frac{\partial \phi}{\partial t}\right)} = \frac{1}{c_s^2}\frac{\partial \phi}{\partial t}\,,$$

and use in the Euler-Lagrange equations:

$$\boxed{\frac{1}{c_s^2}\frac{\partial^2\phi}{\partial t^2} - \frac{\partial^2\phi}{\partial x^2} = 0}$$

You recognize this as the wave equation! The general solution is

$$\phi_R(x - c_s t) + \phi_L(x + c_s t)$$

describing right and left moving waves with speed of propagation c_s (the speed of "sound," hence the subscript s).

Since the notation above is pretty unwieldy, we use, as previously advertised, ∂_t for $\partial/\partial t$, and ∂_x for $\partial/\partial x$ so that, for example, we write the Euler-Lagrange equations as

$$-\partial_t \frac{\partial \mathcal{L}}{\partial_t \phi} - \partial_x \frac{\partial \mathcal{L}}{\partial_x \phi} + \frac{\partial \mathcal{L}}{\partial \phi} = 0\,.$$

Relativistic Fields In the example above we are free to take the speed of light $c = 1$ for the parameter c_s. The solutions to the equation of motion are waves that propagate at the speed of light. Is this then a Lorentz invariant theory? Yes!

We can check this by verifying that if $\phi(x,t)$ is a solution, so is $\phi'(x,t) \equiv \phi(x',t')$ where

$$\begin{aligned} x' &= \gamma(x - \beta t) \\ t' &= \gamma(t - \beta x) \end{aligned} \qquad \gamma = \frac{1}{\sqrt{1-\beta^2}}$$

Alternatively, we can verify that the action integral , $S = \int\int \mathcal{L}\, dt\, dx$, from which the equations are derived, is itself invariant. To this end compute,

$$\begin{aligned}
\frac{\partial\phi'(x,t)}{\partial x} &= \frac{\partial\phi(x',t')}{\partial x} = \frac{\partial\phi(x',t')}{\partial x'}\frac{\partial x'}{\partial x} + \frac{\partial\phi(x',t')}{\partial t'}\frac{\partial t'}{\partial x} = \gamma\frac{\partial\phi(x',t')}{\partial x'} - \beta\gamma\frac{\partial\phi(x',t')}{\partial t'} \\
\frac{\partial\phi'(x,t)}{\partial t} &= \frac{\partial\phi(x',t')}{\partial x} = \frac{\partial\phi(x',t')}{\partial x'}\frac{\partial x'}{\partial t} + \frac{\partial\phi(x',t')}{\partial t'}\frac{\partial t'}{\partial t} = -\beta\gamma\frac{\partial\phi(x',t')}{\partial x'} + \gamma\frac{\partial\phi(x',t')}{\partial t'}
\end{aligned}$$

and use this in the Lagrangian density, (2.1) (with $c_s = c = 1$):

$$\mathcal{L} = \frac{1}{2}\left(\frac{\partial\phi'(x,t)}{\partial t}\right)^2 - \frac{1}{2}\left(\frac{\partial\phi'(x,t)}{\partial x}\right)^2 = \frac{1}{2}\left(\frac{\partial\phi(x',t')}{\partial t'}\right)^2 - \frac{1}{2}\left(\frac{\partial\phi(x',t')}{\partial x'}\right)^2 .$$

Finally integrate this over $\int\int dtdx$ to obtain the action. On the right hand side of the equation change variables of integration , $dx\,dt = \left|\frac{\partial(x,t)}{\partial(x',t')}\right| dx'\,dt' = dx'\,dt'$ to obtain finally

$$\int\int dt\,dx\,\mathcal{L}(\phi'(x,t)) = \int\int dt'\,dx'\,\mathcal{L}(\phi(x',t')) = \int\int dt\,dx\,\mathcal{L}(\phi(x,t))$$

where in the last step we changed the label for the dummy variables of integration from (x', t') to (x, t). This shows invariance of the action integral, $S[\phi'(x,t)] = S[\phi(x,t)]$; the theory is Lorentz invariant.

We could have saved ourselves a lot of time had we taken advantage of the notation designed to exhibit the properties of quantities under Lorentz transformations. We can rewrite the Lagrangian density

$$\mathcal{L} = \frac{1}{2}\left(\frac{\partial\phi(x,t)}{\partial t}\right)^2 - \frac{1}{2}\left(\frac{\partial\phi(x,t)}{\partial x}\right)^2 = \frac{1}{2}\eta^{\mu\nu}\,\partial_\mu\phi\,\partial_\nu\phi = \frac{1}{2}\partial^\mu\phi\partial_\mu\phi$$

As we have seen $\partial_\mu\phi$ transforms as a vector, and the Lagrangian is just the invariant square of this vector!

Klein-Gordon, again While the Lagrangian density above was obtained by a limiting process from a system fo discrete masses and springs, we do not insist in interpreting the relativistic system as some collection of infinitesimal springs and masses. We can take a more general approach to writing a Lagrangian density which may be a good model for some physical system by insisting it written in terms of the appropriate number and type of fields, and constraining it by principles and symmetries we want to incorporate.

For example: Suppose we have a system in 1-spatial dimension that can be described by a single field, $\phi(x,t)$. Moreover, we want it to satisfy an equation of motion of second order (no more than second time derivatives), and we want the action to be invariant under Lorentz transformations. Then the Lagrangian density, $\mathcal{L} = \mathcal{L}(\phi, \partial_\mu\phi)$, can include derivatives only through the invariant $\partial^\mu\phi\partial_\mu\phi$. The simplest Lagrangian density we can think of is the one in the example above, $\mathcal{L} = \frac{1}{2}\partial^\mu\phi\partial_\mu\phi$. The next simplest is

$$\mathcal{L} = \tfrac{1}{2}\partial^\mu\phi\partial_\mu\phi - \tfrac{1}{2}m^2\phi^2.$$

We could have included also a linear term, $g\phi$, with g a a constant, but we can eliminate that term by a field redefinition $\phi \to \phi + g/m^2$. The Euler-Lagrange equation that follows from this Lagrangian density is the 1-spatial dimension version of the Klein-Gordon equation! It is instructive to derive the equation of motion anew, maintaing Lorentz covariance explicitly throughout the computation. We first integrate by parts to recast the action as

$$S[\phi] = \int d^2x\ \left[-\tfrac{1}{2}\phi\partial^2\phi - \tfrac{1}{2}m^2\phi^2\right]$$

where $\partial^2 = \partial^\mu\partial_\mu$. Taking a variation is now trivial,

$$0 = \delta S = -\int d^2x\,\delta\phi\,\left[\partial^2\phi + m^2\phi\right]$$

leading to

$$\left(\partial^2 + m^2\right)\phi(x,t) = 0$$

which you recognize as the Klein-Gordon equation.

2.2 Field Quantization

As we argued in the introduction we need to account for pair creation not just because it is a natural phenomena and because it matters for accuracy, but also because it is required if we are to have a consistent relativistic quantum mechanical theory. We could proceed by enlarging the Hilbert space to include multi-particle states, $|\vec{p}\rangle$, $|\vec{p},\vec{p}'\rangle = |\vec{p}\rangle \otimes |\vec{p}'\rangle$, etc, and then introduce creation/annihilation operators to describe interactions that change particle number. The end result is the same as what we will obtain from tackling head on the problem of quantization of fields.

Before we do so, let's review quantization of classical systems with a discrete set of generalized coordinates q_i, with $i = 1,\dots,N$. We are given a Lagrangian, from which conjugate momenta and a Hamiltonian follow:

$$L = L(q_i,\dot{q}_i) \quad \Rightarrow \quad p_i = \frac{\partial L}{\partial \dot{q}_i} \quad \text{and} \quad H = p_i\dot{q}_i - L$$

Poisson brackets are defined on any functions of p_i and q_i by

$$\{f,g\}_P = \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial q_i}\frac{\partial f}{\partial p_i}. \tag{2.2}$$

Note that here, and in the definition of the Hamiltonian, we have used the generalized Einstein summation convention. One has, in particular, $\{q_i,p_j\}_P = \delta_{ij}$ and $\{q_i,q_j\}_P = 0 = \{p_i,p_j\}_P$. Moreover, the equations of motion in the Hamiltonian formalism can be written as $\dot{p}_j = \{p_j,H\}_P = -\partial H/\partial q_j$ and $\dot{q}_j = \{q_j,H\}_P = \partial H/\partial p_j$. Quantization proceeds by associating an operator on a Hilbert space $\mathcal{H}$ with each of the generalized coordinates and momenta, $q_i \to \hat{q}_i$ and $p_i \to \hat{p}_i$, and replacing the Poisson bracket by ($-i$ times) the commutator of the operators, $\{q_i,p_j\}_P = \delta_{ij} \to -i[\hat{q}_i,\hat{p}_j] = \delta_{ij}$. Similarly $[\hat{q}_i,\hat{q}_j] = 0 = [\hat{p}_i,\hat{p}_j]$. Evolution of operators is given likewise, e.g., $i\dot{\hat{p}}_j = [\hat{p}_j,\hat{H}]$ and $i\dot{\hat{q}}_j = [\hat{q}_j,\hat{H}]$.

We would like to use this same method to quantize field theories. Let's first understand the analogues of conjugate momentum, Hamiltonian and Poisson bracket in classical field theory and only then quantize. Consider the 1-spatial dimensional system of the previous section. How do we take the continuum limit of the Poisson brackets, Eq. (2.2)? It is convenient to start with

$$\sum_j \{q_i,p_j\}_P = \sum_j \delta_{ij} = 1$$

For the continuum limit rewrite $\sum_j p_j = \sum_j \ell(p_j/\ell)$, where we have taken a common separation $\ell_i = \ell$ for simplicity. This suggests $p_j(t) \to \pi(x,t)$, some sort of conjugate momentum density. On the right hand side of the Poisson bracket then $1 = \sum_j \delta_{ij} \to \int dx\, \delta(x - x')$. That is

$$\{\phi(x), \pi(x')\}_P = \delta(x - x')$$

Since

$$\frac{\delta\phi(x)}{\delta\phi(x')} = \delta(x - x')$$

we are led to define

$$\{f, g\}_P = \int dx \left[\frac{\delta f}{\delta\phi(x)}\frac{\delta g}{\delta\pi(x)} - \frac{\delta g}{\delta\phi(x)}\frac{\delta f}{\delta\pi(x)}\right]$$

The momentum conjugate to ϕ can be defined intrinsically (without taking a limit of the discrete system),

$$\pi = \frac{\partial\mathcal{L}}{\partial\dot{\phi}}$$

and the Hamiltonian *density* is defined by

$$\mathcal{H} = \pi\dot{\phi} - \mathcal{L}\,.$$

We are ready to quantize this 1+1 dimensional field theory. We introduce hermitian operators $\hat{\phi}$ and $\hat{\pi}$ on a Hilbert space, and use the quantization prescription that gives us commutation relations from the Poisson brackets,

$$-i[\hat{\phi}(x,t), \hat{\pi}(x',t)] = \delta(x - x')\,, \qquad [\hat{\phi}(x,t), \hat{\phi}(x',t)] = 0 = [\hat{\pi}(x,t), \hat{\pi}(x',t)] \quad (2.3)$$

Note that the commutation relations are given at a common time, but separate space coordinate. The field operators satisfy equations of motion, the Euler-Lagrange equations from the Lagrangian density $\mathcal{L}$. Alternatively, time evolution is given by

$$i\partial_t\hat{\pi}(x,t) = [\hat{\pi}(x,t), \hat{H}] \quad \text{and} \quad i\partial_t\hat{\phi}(x,t) = [\hat{\phi}(x,t), \hat{H}]$$

where the Hamiltonian is

$$\hat{H} = \int dx\, \hat{\mathcal{H}}\,.$$

Let's work this out for the 1+1 Klein-Gordon example:

$$\mathcal{L} = \tfrac{1}{2}\partial^\mu\hat{\phi}\partial_\mu\hat{\phi} - \tfrac{1}{2}m^2\hat{\phi}^2\,.$$

The momentum conjugate to $\hat{\phi}$ is

$$\hat{\pi} = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \partial_t \hat{\phi}$$

and the Hamiltonian density is

$$\hat{\mathcal{H}} = \tfrac{1}{2}\hat{\pi}^2 + \tfrac{1}{2}(\partial_x \hat{\phi})^2 + \tfrac{1}{2}m^2\hat{\phi}^2 .$$

The quantum field equation is just the Klein-Gordon equation,

$$[\partial^2 + m^2]\hat{\phi}(x,t) = 0 .$$

Alternatively,

$$i\partial_t \hat{\pi}(x,t) = [\hat{\pi}(x,t), \hat{H}] = -i(-\partial_x^2\hat{\phi} + m^2\hat{\phi}) \quad \text{and} \quad i\partial_t \hat{\phi}(x,t) = [\hat{\phi}(x,t), \hat{H}] = i\hat{\pi}(x,t)$$

The fields satisfy the equal-time commutation relations (2.3). To understand the content of this theory, we Fourier expand, at fixed time, say $t = 0$

$$\hat{\phi}(x) = \int \frac{dk}{2\pi}\widetilde{\phi}(k)e^{ikx} \qquad \text{and} \qquad \hat{\pi}(x) = \int \frac{dk}{2\pi}\widetilde{\pi}(k)e^{ikx}.$$

That $t = 0$ is implicit here and in the next few lines. Since $\hat{\phi}^\dagger(x) = \hat{\phi}(x)$ we have $\widetilde{\phi}(k)^\dagger = \widetilde{\phi}(-k)$ and $\widetilde{\pi}(k)^\dagger = \widetilde{\pi}(-k)$. The equal-time commutation relations $[\hat{\phi}(x), \hat{\phi}(x')] = 0$ and $[\hat{\pi}(x), \hat{\pi}(x')] = 0$ imply

$$[\widetilde{\phi}(k), \widetilde{\phi}(k')] = 0 = [\widetilde{\pi}(k), \widetilde{\pi}(k')]$$

Then $[\hat{\phi}(x), \hat{\pi}(x')] = i\delta(x - x')$ gives

$$i\delta(x - x') = \int\frac{dk}{2\pi}\int\frac{dk'}{2\pi}\left[\widetilde{\phi}(k)e^{ikx}, \widetilde{\pi}(k')e^{ik'x'}\right] \Rightarrow [\widetilde{\phi}(k), \widetilde{\pi}(k')] = 2\pi i\delta(k + k')$$

The advantage of Fourier transforming shows up first in computing the Hamiltonian, since the ∂_x^2 term is diagonalized:

$$\begin{aligned}\hat{H} &= \int \frac{dk}{2\pi}\left[\tfrac{1}{2}\widetilde{\pi}^\dagger(k)\widetilde{\pi}(k) + \tfrac{1}{2}(k^2 + m^2)\widetilde{\phi}^\dagger(k)\widetilde{\phi}(k)\right] \\ &= \int \frac{dk}{2\pi}\left[\tfrac{1}{2}\widetilde{\pi}^\dagger(k)\widetilde{\pi}(k) + \tfrac{1}{2}\omega_k^2\widetilde{\phi}^\dagger(k)\widetilde{\phi}(k)\right]\end{aligned}$$

I have written ω_k for the energy $\omega_k = E_k = \sqrt{k^2 + m^2}$ for two reasons: (i) we have not shown that k is a momentum so we have no right yet to think of $\sqrt{k^2 + m^2}$ as the energy, and (ii) it becomes clear that the expression for H is that of an infinite sun of linear harmonic oscillators, $\hat{H} = \frac{1}{2}\hat{p}^2 + \frac{1}{2}\omega^2\hat{q}^2$.

Review of QM-simple harmonic oscillator Consider the spring mass system described by

$$L = \tfrac{1}{2}\dot{q}^2 - \tfrac{1}{2}\omega^2 q^2$$

Correspondingly

$$H = \tfrac{1}{2}p^2 + \tfrac{1}{2}\omega^2 q^2$$

Here q and p, aas well as H, are operators on the Hilbert space, but we are suppressing the hat over the symbols since there will be no occasion for confusion:

$$i[p, q] = 1$$

Let

$$\begin{aligned} a &= \frac{1}{\sqrt{2\omega}}(\omega q + ip) \\ a^\dagger &= \frac{1}{\sqrt{2\omega}}(\omega q - ip) \end{aligned}$$

Then $a^\dagger a = 1/2\omega(\omega^2 q^2 + p^2 - i\omega[p, q]) = 1/2\omega(2H - \omega)$ or

$$H = \omega(a^\dagger a + \tfrac{1}{2})$$

Moreover, $[a, a^\dagger] = \frac{1}{2\omega}[\omega q + ip, \omega q - ip] = \frac{1}{2\omega}(2i\omega)[p, q]$ so we have

$$\begin{aligned} [a, a^\dagger] &= 1 \\ [a, a] &= 0 \\ [a^\dagger, a^\dagger] &= 0 \end{aligned}$$

and then

$$\begin{aligned} [H, a^\dagger] &= \omega a^\dagger \\ [H, a] &= -\omega a \end{aligned}$$

We can use these to find the spectrum. Assume that the state $|E\rangle$ is an energy eigenstate:

$$H|E\rangle = E|E\rangle$$

Then

$$H(a^\dagger|E\rangle) = (E + \omega)\left(a^\dagger|E\rangle\right)$$

which means $|E+\omega\rangle = N_+ a^\dagger|E\rangle$ for some normalization constant, N_+. If $|E\rangle$ is normalized, $\langle E|E\rangle = 1$, then

$$\begin{aligned}
1 = \langle E+\omega|E+\omega\rangle &= |N_+|^2\langle E|aa^\dagger|E\rangle \\
&= |N_+|^2\langle E|\left([a,a^\dagger]+a^\dagger a\right)|E\rangle \\
&= |N_+|^2\langle E|\left(1+\frac{1}{2\omega}(2H-\omega)+a^\dagger a\right)|E\rangle \\
&= |N_+|^2\left(\frac{E}{\omega}+\frac{1}{2}\right)
\end{aligned}$$

Similarly, $|E-\omega\rangle = N_- a|E\rangle$ and

$$1 = \langle E-\omega|E-\omega\rangle = |N_-|^2\langle E|a^\dagger a|E\rangle = |N_-|^2\left(\frac{E}{\omega}+\frac{1}{2}\right)$$

So we have an infinite tower of states with energies E, $E\pm\omega$, $E\pm 2\omega, \ldots$ Since the operators $a^\dagger$ and a raise and lower energies, we call them raising and lowering operators, respectively. To avoid a spectrum that is unbounded from below (a catastrophic instability once the system is coupled to external forces), we can insist that for some state $|0\rangle$ the tower stops:

$$a|0\rangle = 0$$

This is the minimum energy state, the "ground state." It has energy $H|0\rangle = \frac{1}{2}\omega|0\rangle$, called the "zero-point" energy. Then $a^\dagger|0\rangle$ has energy $E_1 = \omega + \frac{\omega}{2}$ and so on, $(a^\dagger)^n|0\rangle$ has energy $E_n = \omega(n+\frac{1}{2})$. The tower of states then can be labeled by an integer, $|E_n\rangle = |n\rangle$. We assume they are normalized. Then, from above, $|n+1\rangle = N_+ a^\dagger|n\rangle$ with

$$|N_+|^{-2} = \frac{E_n}{\omega} + \frac{1}{2} = n+1$$

so that

$$|n+1\rangle = \frac{1}{\sqrt{n+1}}a^\dagger|n\rangle = \frac{1}{\sqrt{(n+1)n}}(a^\dagger)^2|n-1\rangle = \cdots = \frac{1}{\sqrt{(n+1)!}}(a^\dagger)^{n+1}|0\rangle$$

Note that since $aa^\dagger = a^\dagger a + 1$ one has $\frac{1}{2}\omega(aa^\dagger + a^\dagger a) = \frac{1}{2}\omega(2a^\dagger a + 1) = H$. This way of writing $H = \frac{1}{2}\omega(aa^\dagger + a^\dagger a)$ hides the zero-point energy.

2.2.1 Particle Interpretation

This suggests introducing

$$\begin{aligned}
a_k &= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\omega_k}}\left(\omega_k\widetilde{\phi}(k) + i\widetilde{\pi}(k)\right) \\
a_k^\dagger &= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\omega_k}}\left(\omega_k\widetilde{\phi}(k)^\dagger - i\widetilde{\pi}(k)^\dagger\right)
\end{aligned} \tag{2.4}$$

These have

$$\begin{aligned}
[a_k, a_{k'}^\dagger] &= \delta(k-k') \\
[a_k, a_{k'}] &= 0 \\
[a_k^\dagger, a_{k'}^\dagger] &= 0
\end{aligned} \tag{2.5}$$

where we have used $\omega_{-k} = \omega_k$. To compute the Hamiltonian, note that

$$a_k^\dagger a_k = \frac{1}{4\pi\omega_k}\left(\omega_k^2 \widetilde{\phi}(k)^\dagger \widetilde{\phi}(k) + \widetilde{\pi}(k)^\dagger \widetilde{\pi}(k) - i\omega_k \widetilde{\pi}(k)^\dagger \widetilde{\phi}(k) + i\omega_k \widetilde{\phi}(k)^\dagger \widetilde{\pi}(k)\right) . \tag{2.6}$$

Since we are going to sum over $\int dk$, we can change variables $k \to -k$ in the last term,

$$\frac{1}{4\pi\omega_k} i\omega_k \widetilde{\phi}^\dagger(k)\widetilde{\pi}(k) \to \frac{i}{4\pi}\widetilde{\phi}^\dagger(-k)\widetilde{\pi}(-k) = \frac{i}{4\pi}\widetilde{\phi}(k)\widetilde{\pi}(-k)$$

The first two terms in (2.6) are the Hamiltonian density and the last two combine into a commutator, so we have

$$\begin{aligned}
\hat{H} &= \frac{1}{2}\int \frac{dk}{2\pi}\left(\widetilde{\pi}(k)^\dagger \widetilde{\pi}(k) + \omega_k^2 \widetilde{\phi}(k)^\dagger \widetilde{\phi}(k)\right) \\
&= \frac{1}{2}\int \frac{dk}{2\pi}\left(4\pi\omega_k a_k^\dagger a_k + \omega_k i[\widetilde{\pi}(k)^\dagger, \widetilde{\phi}(k)]\right) \\
&= \int dk \left(\omega_k a_k^\dagger a_k + \omega_k \delta(0)\right) \\
&= \frac{1}{2}\int dk \omega_k \left(a_k^\dagger a_k + a_k a_k^\dagger\right)
\end{aligned}$$

Let's examine what we have. Assuming that there is a ground state such that $a_k|0\rangle = 0$ for all k, we have a Hilbert space obtained by acting with $a_k^\dagger$'s on $|0\rangle$, e.g.,

$$(a_{k_1}^\dagger)^{n_1}(a_{k_2}^\dagger)^{n_2}\cdots)|0\rangle\, . \tag{2.7}$$

The ground state $|0\rangle$ has energy E_0 given by

$$\hat{H}|0\rangle = \int dk'\,\omega_{k'}\left(a_{k'}^\dagger a_{k'} + \delta(0)\right)|0\rangle = \int dk'\,\omega_{k'}\delta(0)|0\rangle \equiv E_0|0\rangle\, ,$$

and the state $|k\rangle = a_k^\dagger|0\rangle$ has energy

$$\hat{H}|k\rangle = \int dk'\,\omega_{k'}\left(a_{k'}^\dagger a_{k'} + \delta(0)\right)|k\rangle = (\omega_k + E_0)|k\rangle$$

While the zero-point energy, E_0, is infinite, the difference of energy between the state $|k\rangle$ and the ground state is well defined, finite, $\Delta E = \omega_k$. The same is true of the energy of any of the states (2.7). We can only measure energy differences (except in gravitation; that's another story). That is, we are free to add a constant to H without changing the physical content of the theory. So we can redefine

$$\hat{H} = \int dk\, \omega_k a_k^\dagger a_k$$

Examining this more closely write

$$\hat{H} = \frac{1}{2}\int dk\, \omega_k \left(a_k^\dagger a_k + a_k a_k^\dagger - \langle 0|a_k^\dagger a_k + a_k a_k^\dagger|0\rangle \right) = \int dk\, \omega_k a_k^\dagger a_k \,.$$

We say that in the new expression the operators $a_k^\dagger$ and a_k appear "normal ordered" and the operation is called "normal ordering" or "Wick ordering:"

$$:\tfrac{1}{2}(a_k^\dagger a_k + a_k a_k^\dagger): \equiv a_k^\dagger a_k \,.$$

Under $:\xi:$ the operators in ξ commute. The computation above uses the ground state for reference. We will need to make sure that this procedure preserves Lorentz invariance. More on this later.

The energy of the state $|k\rangle$ is $E_k = \omega_k = \sqrt{k^2+m^2}$. So we identify $p = k$ the momentum of the state. This is just as in our introductory presentation of relativistic QM for a single non-interacting particle. This is then interpreted as a single particle state. But now the theory is much richer. For one thing we have many other states, as in (2.7). The Hilbert space of states in (2.7) is called a "Fock space." We interpret them as many particle states. To see this we check a few things:

(i) Energy of $(a_{k_1}^\dagger)^{n_1}(a_{k_2}^\dagger)^{n_2}\cdots|0\rangle$ is $E = n_1 E_{k_1} + n_2 E_{k_2} + \cdots$

(ii) Momentum of $(a_{k_1}^\dagger)^{n_1}(a_{k_2}^\dagger)^{n_2}\cdots|0\rangle$ is $p = n_1 k_1 + n_2 k_2 + \cdots$

The first one follows trivially from the expression for $\hat{H}$ and its action on the states in (2.7). For the second we introduce the operator

$$\hat{p} = \int dk\; k\, a_k^\dagger a_k$$

which gives the desired eigenvalues. We will verify this is the momentum operator below.

From now on we call the operators $a_k^\dagger$ and a_k *creation* and *annihilation* operators, respectively, rather than raising and lowering operators, to remind us that they are adding or taking away a particle from a state. The ground state, $|0\rangle$ is particleless, so we call it the *vacuum state* or just the vacuum.

Statistics As promised, that particles are identical is an automatic consequence of QFT. Note that all particles have the same mass. Moreover, the multiple particle states are automatically symmetric. For example, let $|k_1, k_2\rangle = a^\dagger_{k_1} a^\dagger_{k_2}|0\rangle$. Then

$$|k_1, k_2\rangle = a^\dagger_{k_1} a^\dagger_{k_2}|0\rangle = a^\dagger_{k_2} a^\dagger_{k_1}|0\rangle = |k_2, k_1\rangle$$

where we have used the commutation relations (2.5). More generally $|k_1, \ldots, k_n\rangle$ is symmetric under exchange of any k_i's. This is an unexpected surprise! In NR-QM one simply assumes the wave function is symmetric for bosons (Bose-Einstein statistics), anti-symmetric for fermions (Fermi-Dirac statistics), and it is observed empirically that integer-spin particles are bosons while half-integer spin particles are fermions. There was a hidden assumption in our calculation that resulted in bosonic particles. The assumption is that quantization goes through the replacement $p, q_P \to i[\hat{p}, \hat{q}]_- = i(\hat{p}\hat{q} - \hat{q}\hat{p})$ rather than $p, q_P \to i[\hat{p}, \hat{q}]_+ = i(\hat{p}\hat{q} + \hat{q}\hat{p})$. We will later see that a consistent formulation of QFT requires we use $[\ ,\ ]_-$ for integer spin fields and $[\ ,\ ]_+$ for half-integer spin fields. So not only we will get identical particles automatically, we will get the correct assignment automatically too:

- bosons: spin-0, 1, ...
- fermions: spin-$\frac{1}{2}$, $\frac{3}{2}$, ...

Normalization Note also that

$$\langle k|k'\rangle = \langle 0|a_k a^\dagger_{k'}|0\rangle = \langle 0|[a_k, a^\dagger_{k'}]|0\rangle = \delta(k - k')$$
$$\langle k_1, k_2|k'_1, k'_2\rangle = \langle 0|a_{k_1} a_{k_2} a^\dagger_{k'_1} a^\dagger_{k'_2}|0\rangle = \delta(k_1 - k'_1)\delta(k_2 - k'_2) + \delta(k_1 - k'_2)\delta(k_2 - k'_1)$$

exactly what we expect of identical particle plane wave states. But this is not the desired relativistic normalization. Not a problem, we only need to take for the definition of states

$$|k\rangle = \sqrt{(2\pi)(2E_k)}\,a^\dagger_k|0\rangle \quad \Rightarrow \quad \langle k|k'\rangle = (2\pi)2E_k\delta(k - k')$$

Plane waves are not normalizable states, but we can make normalizable wave packets out of them:

$$|\psi\rangle = \int dk\, \psi(k) a^\dagger_k|0\rangle \;\Rightarrow\; \langle\psi|\psi\rangle = \int dk\, dk'\, \psi(k)^* \psi(k')\langle 0|a_k a^\dagger_{k'}|0\rangle = \int dk\, |\psi(k)|^2 < \infty$$

It is often convenient to define creation and annihilation operators by rescaling the ones we have:

$$\alpha_k = \sqrt{(2\pi)(2E_k)}\,a_k$$

so that $|k\rangle = \alpha_k|0\rangle$ has relativistic normalization. In terms of these

$$\hat{H} = \int (dk)\, E_k \alpha_k^\dagger \alpha_k$$
$$\hat{p} = \int (dk)\, k \alpha_k^\dagger \alpha_k$$

where (dk) is the relativistic invariant measure.

Number Operator The state with n-particles is an energy Eigenstate. It therefore evolves into a state with n particles (itself). Particle number is conserved because there are no interactions (yet). This suggest there must be an observable, that is, a hermitian operator, $\hat{N}$ that

(i) is conserved , $[\hat{N}, \hat{H}] = 0$

(ii) $\langle\psi|\hat{N}|\psi\rangle = N$, the number of particles in state ψ (if it has a definite number of particles)

It should be obvious by now that

$$\hat{N} = \int dk\, a_k^\dagger a_k = \int (dk)\, \alpha_k^\dagger \alpha_k$$

satisfies the above conditions.

We will see later how to generalize these statements to when we include interactions. The startegy will be to derive the form of $\hat{p}^\mu$ and $\hat{N}$ from conserved currents associated with symmetries of $\mathcal{L}$.

Time evolution We have quantized at $t = 0$. In the Heisenberg representation fields have time dependence. So consider $\hat{\phi}(x,t)$, with $\hat{\phi}(x,0)$ corresponding to the field we denoted by $\hat{\phi}(x)$ at $t = 0$ above. Note that the initial choice $t = 0$ is arbitrary since we have time translation invariance (the Lagrangian does not depend explicitly on time). Now,

$$\partial_t \hat{\phi}(x,t) = \hat{\pi}(x,t) = i[\hat{H}, \hat{\phi}(x,t)]\,.$$

The solution is well known,

$$\hat{\phi}(x,t) = e^{iHt}\hat{\phi}(x)e^{-iHt}\,.$$

To understand how this operator acts on the Fock space we cast it in terms of creation and annihilation operators. To this end we invert (2.4)

$$\begin{aligned} \alpha_k &= \omega_k \widetilde{\phi}(k) + i\widetilde{\pi}(k) \\ \alpha_{-k}^\dagger &= \omega_k \widetilde{\phi}(k) - i\widetilde{\pi}(k) \end{aligned} \qquad \Rightarrow \qquad \begin{aligned} \widetilde{\phi}(k) &= \frac{1}{2\omega_k}\left(\alpha_k + \alpha_{-k}^\dagger\right) \\ \widetilde{\pi}(k) &= -\frac{i}{2\omega_k}\left(\alpha_k - \alpha_{-k}^\dagger\right) \end{aligned}$$

Hence

$$\hat{\phi}(x) = \int (dk) \left(\alpha_k e^{ikx} + \alpha_k^\dagger e^{-ikx} \right) .$$

The time dependence is now straightforward:

$$\begin{aligned} \hat{\phi}(x,t) = e^{iHt}\hat{\phi}(x)e^{-iHt} &= \int (dk) \left(\alpha_k e^{-iE_k t + ikx} + \alpha_k^\dagger e^{iE_k t - ikx} \right) \\ &= \int (dk) \left(\alpha_k e^{-ik\cdot x} + \alpha_k^\dagger e^{ik\cdot x} \right) , \end{aligned}$$

where in the last line we have introduced $k^\mu = (E, k)$ and $x^\mu = (t, x)$ to make relativistic invariance explicit. Clearly $\hat{H}$, $\hat{p}$ and $\hat{N}$ are time independent; this is easily seen by taking $\alpha_k \to e^{-iE_k t}\alpha_k$ in the expressions for $\hat{H}$, $\hat{p}$ and $\hat{N}$). Noe the presence of positive and negative energies in $\hat{\phi}(x,t)$. But there are no "negative energy states." Instead there are annihilation operators that subtract honestly positive energies form states, and creation operators that add it.

Note that the field $\hat{\phi}(x,t)$ satisfies the equation of motion,

$$\left(\partial_t^2 - \partial_x^2 + m^2 \right) \hat{\phi}(x,t) = 0$$

This should be the case, as expected from the commutation relations $\partial_t \hat{\pi} = i[\hat{H}, \hat{\pi}]$ and $\partial_t \hat{\phi} = i[\hat{H}, \hat{\phi}]$. But we can verify this directly from the expansion in creation and annihilation operators using

$$(\partial_t^2 - \partial_x^2) e^{-iE_k t + ikx} = -(E_k^2 - k^2) e^{-iE_k t + ikx} = -m^2 e^{-iE_k t + ikx} ,$$

or in relativistic notation,

$$\partial^2 e^{-ik\cdot x} = -k^2 e^{-ik\cdot x} = -m^2 e^{-ik\cdot x} .$$

Momentum Operator We would like the momentum operator to be defined so that $i\hat{p}$ generates translations (and is conserved). We have already defined the operator, so we check now that it does what we want:

$$\begin{aligned} [\hat{p}, \phi(x,t)] &= \int (dk')(dk) \left[k' \alpha_{k'}^\dagger \alpha_{k'}, \alpha_k e^{-iE_k t + ikx} + \alpha_k^\dagger e^{iE_k t - ikx} \right] \\ &= \int (dk) k \left(-\alpha_k e^{-iE_k t + ikx} + \alpha_k^\dagger e^{iE_k t - ikx} \right) \\ &= i\partial_x \int (dk) \left(\alpha_k e^{-iE_k t + ikx} + \alpha_k^\dagger e^{iE_k t - ikx} \right) \\ &= i\partial_x \hat{\phi}(x,t) . \end{aligned}$$

Moreover,

$$[\hat{p}, \hat{H}] = 0$$

so $\hat{p}$ is a constant in time, that is, it is conserved.

Causality While the vanishing of the equal-time commutation relation, $[\phi(x',0),\phi(x,0)] = 0$, was assumed from the start, there is no reason to suspect an analogous result for non-equal times. Let's compute,

$$\begin{aligned}[\phi(x,t),\phi(x',t')] &= \int (dk)(dk') \left[\alpha_k e^{-ik\cdot x} + \alpha_k^\dagger e^{ik\cdot x}, \alpha_{k'} e^{-ik'\cdot x'} + \alpha_{k'}^\dagger e^{ik'\cdot x'}\right] \\ &= \int (dk) \left(e^{ik\cdot(x'-x)} - e^{-ik\cdot(x'-x)}\right) \qquad (2.8)\end{aligned}$$

Notice that the right hand side is explicitly Lorentz invariant and only a function of the difference of 2-vectors, $x'-x$. We define

$$i\Delta(x'-x) = [\phi(x,t),\phi(x',t')]\,.$$

Using Lorentz invariance it is easy to prove that $\Delta(x) = 0$ for space-like x. Since $\Delta(x)$ is Lorentz invariant we can compute it in a boosted frame. For space-like x there is a boost that sets $t = 0$, that is, for space-like separation $x'-x$ there is a frame for which $t' = t$. The commutator vanishes at equal times, and we can then boost back to the original frame to obtain $\Delta(x) = 0$ for $x^2 < 0$. It is not difficult to verify this from the integral above by explicit calculation. One assumes $x^2 < 0$ and continues the integral of the first term in (2.8) much as was done in Fig. 1.1. The second term is continued along a contour on the lower half plane. The two terms cancel each other.

As promised causality is restored. The contribution of the positive and negative energy states cancelled each other. Only, there are no negative energy states. There are annihilation operators.

2.3 $3+1$ Dimensions

Remarkably little changes as we move on to discuss the case of 3 pace and 1 time dimensions. Now

$$L(t) = \int d^3x\,\mathcal{L} = \int d^3x \left[\tfrac{1}{2}(\partial_\mu\phi)^2 - \tfrac{1}{2}m^2\phi^2\right]\,,$$

where the equality is general and the second gives the explicit case of the Lagrangian for Klein-Gordon theory. We have used $\phi = \phi(\vec{x},t) = \phi(x^\mu)$, often also denoted as $\phi(x)$, and $(\partial_\mu\phi)^2 = \eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi$ (sometimes also denoted as $(\partial\phi)^2$, we really like to compress notation). The Poisson brackets are as before, replacing $\delta^{(3)}(\vec{x}'-\vec{x})$ for $\delta(x'-x)$. This goes over directly into the quantum version. So we have equal time commutation relations

$$-i[\hat\phi(\vec{x},t),\hat\pi(\vec{x}',t)] = \delta^{(3)}(\vec{x}-\vec{x}')\,, \qquad [\hat\phi(\vec{x},t),\hat\phi(\vec{x}',t)] = 0 = [\hat\pi(\vec{x},t),\hat\pi(\vec{x}',t)] \qquad (2.9)$$

As before these are solved by

$$\hat{\phi}(\vec{x},0) = \int (dk)\left[\alpha_{\vec{k}} e^{i\vec{k}\cdot\vec{x}} + \alpha^\dagger_{\vec{k}} e^{-i\vec{k}\cdot\vec{x}}\right]$$
$$\hat{\pi}(\vec{x},0) = -i\int (dk) E_{\vec{k}}\left[\alpha_{\vec{k}} e^{i\vec{k}\cdot\vec{x}} - \alpha^\dagger_{\vec{k}} e^{-i\vec{k}\cdot\vec{x}}\right]$$

with

$$[\alpha_{\vec{k}}, \alpha^\dagger_{\vec{k}'}] = (2\pi)^3 2E_{\vec{k}}\delta^{(3)}(\vec{k}-\vec{k}')$$
$$[\alpha_{\vec{k}}, \alpha_{\vec{k}'}] = 0 = [\alpha^\dagger_{\vec{k}}, \alpha^\dagger_{\vec{k}'}]$$

and these are interpreted as annihilation and creation operators for relativistically normalized particle states with mass m: if the vacuum state is $|0\rangle$ then

$$|\vec{k}\rangle = \alpha^\dagger_{\vec{k}}|0\rangle \qquad \text{has} \qquad \langle\vec{k}\,|\vec{k}'\rangle = (2\pi)^3 2E_{\vec{k}}\delta^{(3)}(\vec{k}-\vec{k}')\,.$$

After normal ordering the Hamiltonian is

$$\hat{H} = \int (dk)\, E_{\vec{k}}\, \alpha^\dagger_{\vec{k}}\alpha_{\vec{k}}.$$

The conserved operator

$$\hat{N} = \int (dk)\, \alpha^\dagger_{\vec{k}}\alpha_{\vec{k}}.$$

counts number of particles and

$$\hat{\vec{p}} = \int (dk)\, \vec{k}\, \alpha^\dagger_{\vec{k}}\alpha_{\vec{k}}.$$

are the conserved momentum operators and generate translations. As before the particles are identical and multiplarticle states satisfy Bose-Einstein statistics. In contrast to the $1+1$ case, in 3-spatial dimensions we can speak meaningfully of the spin of a particle. It must correspond to a quantum number that transforms under rotations. Our field is invariant under Lorentz transformations, $\phi(x) \to \phi'(x) = \phi(x')$, where $x' = \Lambda x$, and this is results in spinless particles. The spin-statistics connection comes out automatically: spin-0 identical particles satisfy Bose-Einstein statistics.

Time evolution is still given by

$$\begin{aligned}\hat{\phi}(x) = \hat{\phi}(\vec{x},t) &= e^{iHt}\phi(\vec{x},0)e^{-iHt}\\ &= \int (dk)\left[\alpha_{\vec{k}} e^{i\vec{k}\cdot\vec{x}-iE_{\vec{k}}t} + \alpha^\dagger_{\vec{k}} e^{-i\vec{k}\cdot\vec{x}+iE_{\vec{k}}t}\right]\\ &= \int (dk)\left[\alpha_{\vec{k}} e^{-ik\cdot x} + \alpha^\dagger_{\vec{k}} e^{ik\cdot x}\right]\end{aligned}$$

where $k^0 = E_{\vec{k}}$. The operator $\hat{\phi}(x)$ satisfies the Klein-Gordon equation,

$$\left(\partial^2 + m^2\right)\hat{\phi}(x) = 0$$

which is the Euler-Lagrange equation for the Lagrangian density given above.

We will encounter later the product $\hat{\phi}(x_1)\hat{\phi}(x_2)$, and we will need its relation to the normal ordered product $:\hat{\phi}(x_1)\hat{\phi}(x_2):$. It is convenient to introduce "positive and negative frequency operators,"

$$\begin{aligned}
\hat{\phi}^{(-)}(x) &= \int (dk) e^{ik\cdot x}\alpha^\dagger_{\vec{k}}\,, \\
\hat{\phi}^{(+)}(x) &= \int (dk) e^{-ik\cdot x}\alpha_{\vec{k}}\,.
\end{aligned}$$

Then

$$\begin{aligned}
\hat{\phi}(x_1)\hat{\phi}(x_2) &= \left(\hat{\phi}^{(+)}(x_1) + \hat{\phi}^{(-)}(x_1)\right)\left(\hat{\phi}^{(+)}(x_2) + \hat{\phi}^{(-)}(x_2)\right) \\
&= \hat{\phi}^{(+)}(x_1)\hat{\phi}^{(+)}(x_2) + \hat{\phi}^{(+)}(x_1)\hat{\phi}^{(-)}(x_2) + \hat{\phi}^{(-)}(x_1)\hat{\phi}^{(+)}(x_2) + \hat{\phi}^{(-)}(x_1)\hat{\phi}^{(-)}(x_2) \\
&= \hat{\phi}^{(+)}(x_1)\hat{\phi}^{(+)}(x_2) + \hat{\phi}^{(-)}(x_2)\hat{\phi}^{(+)}(x_1) + \hat{\phi}^{(-)}(x_1)\hat{\phi}^{(+)}(x_2) \\
&\qquad + \hat{\phi}^{(-)}(x_1)\hat{\phi}^{(-)}(x_2) + [\hat{\phi}^{(+)}(x_1), \hat{\phi}^{(-)}(x_2)] \\
&= :\hat{\phi}(x_1)\hat{\phi}(x_2): + [\hat{\phi}^{(+)}(x_1), \hat{\phi}^{(-)}(x_2)]
\end{aligned}$$

Hence the difference between the product and the normal ordered product is a c-number,

$$\begin{aligned}
\Delta_+(x_2 - x_1) &\equiv [\hat{\phi}^{(+)}(x_1), \hat{\phi}^{(-)}(x_2)] \\
&= \int (dk_1)(dk_2)\, e^{ik_1\cdot x_1 - ik_2\cdot x_2}[\alpha_{\vec{k}_1}, \alpha^\dagger_{\vec{k}_2}] \\
&= \int (dk)\, e^{-ik\cdot(x_2 - x_1)}\,. \qquad (2.10)
\end{aligned}$$

That Δ_+ is only a function of the difference is the result of the explicit calculation above. You may recognize this as the integral in (1.1). Note that

$$\left(\partial^2 + m^2\right)\Delta_+(x) = 0 \qquad \text{for } x \neq 0,$$

so $\Delta_+(x)$ is a solution of the Klein-Gordon equation that does not vanish for spacelike argument.

Quantum theory is weird, of course, but QFT is even weirder. Consider this. The expectation value of $\hat{\phi}(x)$ in the vacuum state is zero at any point x,

$$\langle 0|\hat{\phi}(x)|0\rangle = 0$$

But the expectation value of the square $\hat{\phi}^2(x)$ is infinite:

$$\langle 0|\hat{\phi}^2(x)|0\rangle = \Delta_+(0) = \int (dk)\,.$$

Fluctuations of quantum fields at any point are wild, even for the simplest empty state! The problem arises from localization: we should not insist in determining the field precisely in an arbitrarily small region of space (in this case, one point). In homework you will show that the square remains finite for the field smeared over a region.

Chapter 3

Symmetries

A famous theorem of Wigner shows that symmetries in a quantum theory must correspond to either unitary or anti-unitary operators. It seems fit to start with a review of what is meant by this. We will then proceed to study continuous symmetries, all represented by unitary operators. We will then turn our attention to discrete symmetries. It is then, in presenting time-reversal symmetry, that we will encounter anti-unitary operators.

3.1 Review of unitary and anti-unitary operators

The bra/ket notation is not quite suitable for anti-linear operators. So for this section we use the following notation:

- States are denoted by wave-functions: $\psi, \chi, \ldots$
- c-numbers are lowercase latin letters: $a, b, \ldots$
- Operators are uppercase: $A, B, \ldots, U, V, \ldots, \Omega, \ldots$
- Inner product and norm: (ψ, χ) and $\|\psi\|^2 = (\psi, \psi)$

An operator A is *linear* if $A(a\psi + b\chi) = aA\psi + bA\chi$. The *hermitian conjugate* $A^\dagger$ of an operator A, is such that for all χ, ψ

$$(\chi, A^\dagger\psi) = (A\chi, \psi)\,.$$

This is consistent only if A is linear:

$$\begin{aligned}
(A(a\psi + b\chi), \rho) &= (a\psi + b\chi, A^\dagger\rho) \\
&= a^*(\psi, A^\dagger\rho) + b^*(\chi, A^\dagger\rho) \\
&= a^*(A\psi, \rho) + b^*(A\chi, \rho) \\
&= (aA\psi + bA\chi, \rho)
\end{aligned}$$

for any ρ.

An invertible operator U is *unitary* if

$$(U\psi, U\chi) = (\psi, \chi) \qquad \text{(and therefore } \|U\psi\| = \|\psi\|\text{)}.$$

Unitary operators are linear. Proof:

$$\begin{aligned}
\|U(a\psi + b\chi) - aU\psi - bU\chi\|^2 &= \|U(a\psi + b\chi)\|^2 - 2\mathrm{Re}\,[a(U(a\psi + b\chi), U\psi)] \\
&\quad - 2\mathrm{Re}\,[b(U(a\psi + b\chi), U\chi)] + |a|^2 \|U\psi\|^2 + |b|^2 \|U\chi\|^2 \\
&= \|a\psi + b\chi\|^2 - 2\mathrm{Re}\,[a(a\psi + b\chi, \psi)] \\
&\quad - 2\mathrm{Re}\,[b(a\psi + b\chi, \chi)] + |a|^2 \|\psi\|^2 + |b|^2 \|\chi\|^2 \\
&= 0
\end{aligned}$$

where we have used unitarity in going form the first to the second line and we have expanded all terms in going from the second to the third line. Since only zero has zero norm we have $U(a\psi + b\chi) - aU\psi - bU = 0$, completing the proof. The inverse of a unitary operator is its hermitian conjugate:

$$(\psi, U^{-1}\chi) = (U\psi, U(U^{-1}\chi)) = (U\psi, \chi) = (\psi, U^\dagger\chi)$$

for any ψ, χ.

An invertible operator Ω is *anti-unitary* if for all ψ, χ

$$(\Omega\psi, \Omega\chi) = (\chi, \psi) \qquad \text{(notice inverted order).}$$

An operator $\rotatebox{90}{A}$ is *anti-linear* if for all ψ, χ

$$\rotatebox{90}{A}(a\psi + b\chi) = a^*\rotatebox{90}{A}\psi + b^*\rotatebox{90}{A}\chi$$

Anti-unitary operators are anti-linear. The proof is the same as in linearity of unitary operators, $\|\Omega(a\psi + b\chi) - a^*\Omega\psi + b^*\Omega\chi\|^2 = 0$. Example: the complex conjugation operator, Ω_c. Clearly

$$\Omega_c(a\psi + b\chi) = a^*\Omega_c\psi + b^*\Omega_c\chi \qquad \text{and} \qquad (\Omega_c\psi, \Omega_c\chi) = (\chi, \psi)\,.$$

You can show that the products U_1U_2 and $\Omega_1\Omega_2$ of two unitary or two anti-unitary operators are unitary operators, while the products $U\Omega$ and ΩU of a unitary and an anti-unitary operators are anti-unitary.

Symmetries What properties are required of operator symmetries in QM? If A is an operator on the Hilbert space, $\mathcal{F} = \{\psi\}$, then it is a symmetry transformation if it preserves probabilities, $|(A\psi, A\chi)|^2 = |(\psi, \chi)|^2$. Wigner showed that the only two possibilities are A is unitary or anti-unitary.

Symmetry transformations as action on operators: for unitary operators, the symmetry transformation $\psi \to U\psi$, for all states, gives

$$(\psi, A\chi) \to (U\psi, AU\chi) = (\psi, U^\dagger AU\chi)$$

So we can transform instead operators, via $A \to U^\dagger AU$. For anti-unitary operators the hermitian conjugate is not defined, so we do not have "$\Omega^\dagger = \Omega^{-1}$". But

$$(\psi, A\chi) \to (\Omega\psi, A\Omega\chi) = (\Omega^{-1}A\Omega\chi, \Omega^{-1}\Omega\psi) = (\psi, (\Omega^{-1}A\Omega)^\dagger\chi)\,,$$

which is not very useful. For expectation values of observables, $A^\dagger = A$,

$$(\psi, A\psi) \to (\Omega\psi, A\Omega\psi) = (A\Omega\psi, \Omega\psi) = (\psi, \Omega^{-1}A\Omega\psi)$$

and in this limited sense, $A \to \Omega^{-1}A\Omega$.

3.2 Continuous symmetries, Generators

Consider a family of unitary transformations, $U(s)$, where s is a real number indexing the unitary operators. We assume $U(0) = \mathbb{1}$, the identity operator. Furthermore, assume $U(s)$ is continuous, differentiable. Then expanding about zero,

$$U(\epsilon) = \mathbb{1} + i\epsilon T + \mathcal{O}(\epsilon^2) \tag{3.1}$$

$$U(\epsilon)^\dagger U(\epsilon) = \mathbb{1} = (\mathbb{1} - i\epsilon T^\dagger)(\mathbb{1} + i\epsilon T) + \mathcal{O}(\epsilon^2) \quad \Rightarrow \quad T^\dagger = T \tag{3.2}$$

$$iT \equiv \left.\frac{dU(s)}{ds}\right|_{s=0} \tag{3.3}$$

The operator T is called a symmetry *generator*.

Now, the product of N transformations is a transformation, so consider

$$\left(\mathbb{1} + i\frac{s}{N}T\right)^N$$

In the limit $N \to \infty$, $\left(\mathbb{1} + i\frac{s}{N}T\right)$ is unitary , and therefore so is $\left(\mathbb{1} + i\frac{s}{N}T\right)^N$. This is a vulgarized version of the exponential map,

$$U(s) = \lim_{N\to\infty}\left(\mathbb{1} + i\frac{s}{N}T\right)^N = e^{isT}\,.$$

Note that

$$U(s)^\dagger U(s) = e^{-isT}e^{isT} = \mathbb{1}\,,$$

as it should. Also, $A \to U(s)^\dagger AU(s)$ becomes, for $s = \epsilon$ infinitesimal,

$$A \to A + i\epsilon[A, T] \qquad \text{or} \qquad \delta A = i\epsilon[A, T]\,.$$

A symmetry of the Hamiltonian has

$$U(s)^\dagger H U(s) = H \quad \Rightarrow \quad [H,T] = 0\,.$$

Since for any operator (in the Heisenberg picture)

$$i\frac{dA}{dt} = [H,A] + i\frac{\partial A}{\partial t}$$

we have that a symmetry generator T is a constant, $dT/dt = 0$ ($T(t)$ has no explicit time dependence). Conversely a constant hermitian operator T defines a symmetry:

$$dT/dt = 0 \quad \Rightarrow \quad [H,T] = 0 \quad \Rightarrow e^{-isT} H e^{isT} = H\,.$$

Let's connect this to symmetries of a Lagrangian density (or, more precisely, of the action integral).

3.3 Noether's Theorem

Consider a Lagrangian $L(t) = \int d^3x\, \mathcal{L}(\phi^a, \partial_\mu \phi^a)$, where the Lagrangian density is a function of N real fields ϕ^a, $a = 1, \ldots, N$. We investigate the effect of a putative symmetry transformation

$$\phi^a \to \phi'^a = \phi^a + \delta\phi^a, \quad \text{where } \delta\phi^a = \epsilon D^{ab}\phi^b,$$

with ϵ an infinitesimal parameter and $(D\phi)^a = D^{ab}\phi^b$ is a linear operator on the collection of fields (and may contain derivatives). Consider then $\mathcal{L}(\phi'^a, \partial_\mu \phi'^a)$. Suppose that by explicit computation we find

$$\delta\mathcal{L} = \mathcal{L}(\phi') - \mathcal{L}(\phi) = \epsilon \partial_\mu \mathcal{F}^\mu\,,$$

that is, that the variation of the Lagrangian density vanishes up to a derivative, the divergence of a four-vector. We do not require that the variation of $\mathcal{L}$ vanishes since we want invariance of the action integral, to which total derivatives contribute only vanishing surface terms. For any variation $\delta\phi$, not necessarily of the form displayed above,

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\phi^a}\delta\phi^a + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^a)}\partial_\mu\delta\phi^a\,.$$

If ϕ^a are solutions to the equations of motion, the first term can be rewritten and the two terms combined,

$$\delta\mathcal{L} = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^a)}\right)\delta\phi^a + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^a)}\partial_\mu\delta\phi^a = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^a)}\delta\phi^a\right)\,.$$

Using the variation that by assumption vanishes up to a total derivative we have

$$\partial_\mu \mathcal{F}^\mu = \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^a)} D^{ab} \phi^b \right)$$

so that

$$\boxed{J^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^a)} D^{ab} \phi^b - \mathcal{F}^\mu \quad \text{satisfies} \quad \partial_\mu J^\mu = 0\,.}$$

This is Noether's theorem. Note that this can be written in terms of a generalized momentum conjugate, $J_\mu = \pi^a_\mu D^{ab} \phi^b - \mathcal{F}_\mu$. Since J^μ is a conserved current, the spatial integral of the time component is a conserved "charge,"

$$T = \int d^3x\, J^0 \qquad \text{has} \quad \frac{dT}{dt} = 0\,.$$

The conserved charge is nothing but the generator of a continuous symmetry. We can show from its definition that it commutes with the Hamiltonian.

Let's look at some examples:

Translations The transformation is induced by $x^\mu \to x^\mu + \epsilon a^\mu$. Clearly

$$\delta \mathcal{L} = \epsilon a^\mu \partial_\mu \mathcal{L} \quad \Rightarrow \quad \mathcal{F}^\mu = a^\mu \mathcal{L}\,.$$

But $\mathcal{L} = \mathcal{L}(x)$, is a function of x^μ only through its dependence on $\phi^a(x)$, so

$$\delta \phi^a = \epsilon a^\mu \partial_\mu \phi^a \quad \Rightarrow \quad J^\mu_{(a)} = \pi^{a\mu} a^\nu \partial_\nu \phi^a - a^\mu \mathcal{L} = a_\nu (\pi^{a\mu} \partial^\nu \phi^a - \eta^{\mu\nu} \mathcal{L})\,.$$

Since a^ν is arbitrary, we can choose it to be alternatively along the direction of any of the four independent directions of space-time, and we then have in fact four independently conserved currents:

$$\boxed{T^{\mu\nu} = \pi^{a\mu} \partial^\nu \phi^a - \eta^{\mu\nu} \mathcal{L}}$$

This is the *energy and momentum tensor*, also known as the *stress-energy tensor*. The first index refers to the conserved current and the second labels which of the four currents.

The conserved "charges" are

$$P^\mu = \int d^3x\, T^{0\mu}\,.$$

They are associated with the transformation $x \to x + a$ and hence they are momenta. For example,

$$P^0 = \int d^3x (\pi^a \phi^a - \mathcal{L}) = \int d^3x\, \mathcal{H} = H$$

as expected.

Example: In our Klein-Gordon field theory

$$\mathcal{L} = \tfrac{1}{2}(\partial_\mu\phi)^2 - \tfrac{1}{2}m^2\phi^2 \tag{3.4}$$

$$\pi^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} = \partial^\mu\phi \tag{3.5}$$

$$\Rightarrow \quad T^{\mu\nu} = \partial^\mu\phi\partial^\nu\phi - \eta^{\mu\nu}\mathcal{L} \tag{3.6}$$

We can verify this is conserved, by using equations of motion,

$$\begin{aligned}\partial_\mu T^{\mu\nu} &= \partial^2\phi\partial^\nu\phi + \partial^\mu\phi\partial^\nu\partial_\mu\phi - \partial^\nu(\tfrac{1}{2}(\partial_\mu\phi)^2 - \tfrac{1}{2}m^2\phi^2)\\ &= -m^2\phi\partial^\nu\phi + \partial^\mu\phi\partial^\nu\partial_\mu\phi - \partial^\mu\phi\partial^\nu\partial_\mu\phi + m^2\phi\partial^\nu\phi\\ &= 0\end{aligned}$$

Turning to the case of quantum filed theory compute now the conserved 4-momentum operator in terms of creation and annihilation operators. For the temporal component, $P^0 = H$, the computation was done in the previous chapter. For the spatial components we have

$$P^i = \int d^3x\, T^{0i} = \int d^3x\,(\pi\partial^i\phi - \eta^{0i}\mathcal{L}) = \int d^3x\,\pi\partial^i\phi$$

so that

$$\begin{aligned}P^i &= \int d^3x\,\partial_t\phi\partial^i\phi\\ &= \int d^3x \int (dk')(dk)\,[(-iE_{\vec{k}'})(\alpha_{\vec{k}'}e^{-ik'\cdot x} - \alpha^\dagger_{\vec{k}'}e^{ik'\cdot x})][(-ik^i)(\alpha_{\vec{k}}e^{-ik\cdot x} - \alpha^\dagger_{\vec{k}}e^{ik\cdot x})]\\ &= \int (dk')(dk)\,(-iE_{\vec{k}'})(-ik^i)\left[\left((2\pi)^3\delta^{(3)}(\vec{k}'+\vec{k})\alpha_{\vec{k}'}\alpha_{\vec{k}}e^{-i(E_{\vec{k}'}+E_{\vec{k}})t} + \text{h.c.}\right)\right.\\ &\qquad\left. - \left((2\pi)^3\delta^{(3)}(\vec{k}'-\vec{k})\alpha_{\vec{k}'}\alpha^\dagger_{\vec{k}}e^{-i(E_{\vec{k}'}-E_{\vec{k}})t} + \text{h.c.}\right)\right]\\ &= \frac{1}{2}\int (dk)k^i\left[\alpha_{\vec{k}}\alpha_{-\vec{k}}e^{-2iE_{\vec{k}}t} + \alpha^\dagger_{\vec{k}}\alpha^\dagger_{-\vec{k}}e^{2iE_{\vec{k}}t} + \alpha^\dagger_{\vec{k}}\alpha_{\vec{k}} + \alpha_{\vec{k}}\alpha^\dagger_{\vec{k}}\right]\end{aligned}$$

In the last line the first two terms are odd under $\vec{k} \to -\vec{k}$ so they vanish upon integration and we finally have

$$P^i = \int (dk)\,\tfrac{1}{2}k^i\left(\alpha^\dagger_{\vec{k}}\alpha_{\vec{k}} + \alpha_{\vec{k}}\alpha^\dagger_{\vec{k}}\right) = \int (dk)k^i\alpha^\dagger_{\vec{k}}\alpha_{\vec{k}}\,.$$

It follows that

$$[\vec{P}, \alpha^\dagger_{\vec{k}}] = \vec{k}\alpha^\dagger_{\vec{k}} \quad \text{so that} \quad P^\mu|\vec{k}\rangle = k^\mu|\vec{k}\rangle\,.$$

More generally, we compute the commutator of P^μ and $\phi(\vec{y},t)$. Since $P\mu$ is time independent we can always choose to compute at equal times. Then the equal time commutator is fixed by the commutation relations from canonical quantization:

$$[P^i, phi(x)] = \int d^3x\, [\pi\partial^i\phi(\vec{x},t), \phi(\vec{y},t] = -i\partial^i\phi(\vec{y},t)\,.$$

The case for P^0 is a bit more involved, but we know $P^0 = H$ so we know the result without need to compute. It follows that

$$U(a)\phi(x)U(a)^\dagger = e^{ia\cdot P}\phi(x)e^{-ia\cdot P} = \phi(x) + ia_\mu[P^\mu, \phi(x)] + \cdots = \phi(x+a)$$

or $\phi'(x) = U(a)^\dagger\phi(x)U(a) = \phi(x-a)$.

Lorentz Transformations The transformation of states was introduced earlier, $U(\Lambda)|\vec{p}\rangle = |\Lambda\vec{p}\rangle$. Assuming the vacuum state is invariant under Lorentz transformations we then may take $U(\Lambda)\alpha^\dagger_{\vec{p}}U(\Lambda)^\dagger = \alpha^\dagger_{\Lambda\vec{p}}$ and $U(\Lambda)\alpha_{\vec{p}}U(\Lambda)^\dagger = \alpha_{\Lambda\vec{p}}$. Therefore

$$U(\Lambda)\phi(x)U(\Lambda)^\dagger = \phi(\Lambda x) \quad \text{or} \quad U(\Lambda)^\dagger\phi(x)U(\Lambda) = \phi(\Lambda^{-1}x)$$

Let $\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \epsilon\omega^\mu{}_\nu$ with ϵ infinitesimal. Then $\omega^{\mu\nu} = -\omega^{\nu\mu}$, and therefore has $4\times 3/2 = 6$ independent components, three for rotations (ω^{ij}) and three for boosts (ω^{0i}). We assume $\mathcal{L} = \mathcal{L}(\phi, \partial_\mu\phi)$ is Lorentz invariant. What does this mean? For scalar fields $\phi(x) \to \phi'(x) = \phi(\Lambda^{-1}x) = \phi(x')$ and by the chain rule $\partial_\mu\phi(x') = \partial_\mu x'^\lambda\partial'_\lambda\phi(x') = (\Lambda^{-1})^\lambda{}_\mu\partial'_\lambda\phi(x') = \Lambda_\mu{}^\lambda\partial'_\lambda\phi(x')$, which gives

$$\mathcal{L}(\phi, \partial_\mu\phi) \to \mathcal{L}' = \mathcal{L}(\phi'(x), \partial_\mu\phi'(x)) = \mathcal{L}(\phi(x'), \partial_\mu\phi(x')) = \mathcal{L}(\phi(x'), \Lambda_\mu{}^\lambda\partial'_\lambda\phi(x'))\,.$$

Invariance means $\mathcal{L}'$ has the same functional dependence as $\mathcal{L}$ but in terms of x', that is, $\mathcal{L}$ is a scalar, $\mathcal{L}(x) \to \mathcal{L}'(x) = \mathcal{L}(x')$. This gives

$$\mathcal{L}(\phi(x'), \Lambda_\mu{}^\lambda\partial'_\lambda\phi(x')) = \mathcal{L}(\phi(x'), \partial'_\mu\phi(x'))\,.$$

For example, $\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi \to \eta^{\mu\nu}\Lambda_\mu{}^\lambda\Lambda_\nu{}^\sigma\partial'_\lambda\phi\partial'_\sigma\phi = \eta^{\lambda\sigma}\partial'_\lambda\phi\partial'_\sigma\phi$ works, but not so $a^\mu\partial_\mu\phi$ for constant vector a^μ since a^μ does not transform (it is a fixed constant).

We assume $\mathcal{L}$ is Lorentz invariant and compute:

$$\begin{aligned}
\delta\phi &= \phi(x^\mu - \epsilon\omega^\mu{}_\nu x^\nu) - \phi(x^\mu) = -\epsilon\omega^\mu{}_\nu x^\nu\partial_\mu\phi = -\epsilon\omega^{\mu\nu}x_\nu\partial_\mu\phi \\
\delta\mathcal{L} &= -\epsilon\omega^{\mu\nu}x_\nu\partial_\mu\mathcal{L} = -\left[\partial_\mu(\epsilon\omega^{\mu\nu}x_\nu\mathcal{L}) - \epsilon\omega^{\mu\nu}\eta_{\mu\nu}\mathcal{L}\right] = -\partial_\mu(\epsilon\omega^{\mu\nu}x_\nu\mathcal{L})
\end{aligned}$$

and then the conserved currents are

$$\begin{aligned}
J^\mu_{(\omega)} &= -\pi^\mu\omega^{\lambda\sigma}x_\sigma\partial_\lambda\phi + \omega^{\mu\nu}x_\nu\mathcal{L} \\
&= -\omega^{\lambda\sigma}(\pi^\mu x_\sigma\partial_\lambda\phi - \delta^\mu{}_\lambda x_\sigma\mathcal{L})
\end{aligned}$$

or, since $\omega^{\mu\nu}$ is arbitrary, we have six conserved currents,

$$\begin{aligned} M^{\mu\nu\lambda} &= (\pi^\mu x^\nu \partial^\lambda - \eta^{\mu\lambda} x^\nu \mathcal{L}) - \nu \leftrightarrow \lambda \\ &= x^\nu T^{\mu\lambda} - x^\lambda T^{\mu\nu} \,. \end{aligned}$$

This has

$$M^{\nu\lambda} = \int d^3x \, M^{0\nu\lambda}$$

as generator of rotations (M^{ij}) and boosts (M^{0i}).

Comments:

(i) The expression for $M^{\mu\nu\lambda}$ is specific for scalar fields since we have used $U^\dagger(\Lambda)\phi(x)U(\Lambda) = \phi(\Lambda^{-1}x)$. This *defines* scalar fields. We expect, for example, that a vector field, $A^\mu(x)$ will transform like $\partial^\mu\phi(x)$,

$$U^\dagger(\Lambda)A_\mu(x)U(\Lambda) = \Lambda_\mu{}^\lambda A_\lambda(\Lambda^{-1}x) \quad \text{or} \quad U(\Lambda)^\dagger A^\mu(x)U(\Lambda) = \Lambda^\mu{}_\lambda A^\lambda(\Lambda^{-1}x)$$

Then

$$\begin{aligned} \delta A^\mu &= (\delta^\mu{}_\nu + \epsilon\omega^\mu{}_\nu)A(x^\lambda - \epsilon\omega^\lambda{}_\sigma x^\sigma) - A^\mu(x) \\ &= -\epsilon\big[\underbrace{\omega^\lambda{}_\sigma x^\sigma \partial_\lambda A^\mu}_{\text{as before}} - \underbrace{\omega^\mu{}_\nu A^\nu}_{\text{new term}}\big] \end{aligned}$$

This then gives a new term in $J^\mu_{(\omega)}$ of the form $\frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\rho)}(\omega^{\rho\nu}A_\nu)$ leading to an additional term in $M^{\mu\rho\nu}$,

$$\begin{aligned} \Delta M^{\mu\rho\nu} &= \frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\rho)}A^\nu - \frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\nu)}A^\rho \\ &= \frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\lambda)}A^\sigma(\delta^\rho_\lambda\delta^\nu_\sigma - \delta^\rho_\sigma\delta^\nu_\lambda) \end{aligned} \tag{3.7}$$

There is a generalization of the matrices $(\mathcal{I}^{\rho\nu})_{\lambda\sigma} = \delta^\rho_\lambda\delta^\nu_\sigma - \delta^\rho_\sigma\delta^\nu_\lambda$ to the case of fields other than vectors, that is, fields that have other Lorentz transformations, like spin$\frac{1}{2}$ fields. The generalization has $(\mathcal{I}^{\rho\nu})_{\lambda\sigma} \to (\mathcal{I}^{\rho\nu})_{ab}$, where $\rho\nu$ labels the matrices and ab give the specific matrix elements with a, b running over the number of components of the new type of field. The matrices $\mathcal{I}^{\rho\nu}$ satisfy the same commutation relations as $M^{\rho\nu}$. We will explore this in more detail later. For now, the important point is that physically we have a clear interpretation:

$$M^{\mu ij} = \underbrace{x^i T^{\mu j} - x^j T^{\mu i}}_{\text{orbital ang mom } \epsilon^{ijk}L^k} + \underbrace{\Delta M^{\mu ij}}_{\text{intrinsic ang mom (spin) } \epsilon^{ijk}S^k}$$

(ii) The conserved quantities M^{ij} are angular momentum. What are M^{0i}? To get some understanding consider a classical "field" for point particles, with

$$T^{0i} = \sum_n p_n^i \delta^{(3)}(\vec{x} - \vec{x}_m(t)) \quad \text{so that} \quad P^i = \sum_n p_n^i \,.$$

Then we can compute

$$M^{ij} = \int d^3x \, (x^j T^{0i} - x^i T^{0j}) = \sum_n (x_n^i p_n^j - x_n^j p_n^i) = \epsilon^{ijk} \sum_n L_n^k$$

as expected. Turning to the mysterious components,

$$M^{0i} = \int d^3x \, (x^0 T^{0i} - x^i T^{00}) = x^0 P^i - \int d^3x \, x^i T^{00}$$

We can now see what conservation of M^{0i}, namely $\frac{dM^{0i}}{dt} = 0$, gives:

$$0 = P^i - \frac{d}{dt} \int d^3x \, x^i T^{00}$$

The quantity $\int d^3x \, x^i T^{00}$ is a relativistic generalization of center of mass, say, a "center of energy." This can be seen from T^{00} being an energy density, so that when all particles are at rest it corresponds to the mass density. So conservation of M^{0i} means that the center of energy motion is given by the total momentum. This is the relativistic analogue of $\vec{P} = M\vec{V}$ where $\vec{P}$ is the total momentum of the system, M its total mass and $\vec{V}$ the velocity of the center of mass. In the relativistic case we have that

$$\frac{P^i}{P^0} = \text{velocity of "C.M."} = \frac{\frac{d}{dt} \int d^3x \, x^i T^{00}}{P^0}$$

That the "C.M." moves with a constant speed is a relativistic conservation law.

3.4 Internal Symmetries

We have discussed symmetries that have to be present in any of the models we will care about: invariance under translations and under Lorentz transformations, or *Poincare invariance*. These symmetries transform fields at one space-time point to fields at other space-time points. In this sense they are geometrical. But there may be, in addition, symmetries that transform fields at the same space-time point. These are not generic, they are specific to each model. They are called *internal symmetries*. The name comes from the associated conserved quantities giving "internal" characteristics fo the particles. For example, baryon number or isotopic spin are symmetries and they are associated with the baryon number or the isotopic spin of particles.

Let's study a simple example. Consider a model with two real scalar fields,

$$\mathcal{L} = \sum_{n=1}^{2} \tfrac{1}{2}\partial^\mu \phi_n \partial_\mu \phi_n - V(\sum_{n=1}^{2} \phi_n^2)$$

This satisfies $\mathcal{L}(\phi'_n) = \mathcal{L}(\phi_n)$ where

$$\phi'_1(x) = \cos\theta \phi_1(x) - \sin\theta \phi_2(x)$$
$$\phi'_2(x) = \sin\theta \phi_1(x) + \cos\theta \phi_2(x)$$

or more concisely

$$\vec{\phi}' = R\vec{\phi}, \quad \text{where} \quad \vec{\phi} = \begin{pmatrix} \phi_1(x) \\ \phi_2(x) \end{pmatrix} \text{ and } R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{3.8}$$

so that invariance is just the statement that the length of a two dimensional vector is invariant under rotations, $(R\vec{r}) \cdot (R\vec{r}) = \vec{r} \cdot \vec{r}$. Note that $R^T R = \mathbb{1}$. The set of symmetry transformations form a group, $O(2)$; the rotations given explicitly in (3.8) have determinant +1 and tehy form the group $SO(2)$. For this discussion we focus on transformations that can be reached from $\mathbb{1}$ continuously, so we restrict our attention to $SO(2)$. Setting $\theta = \epsilon$ infinitesimal in (3.8),

$$R = \mathbb{1} + \epsilon \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \mathbb{1} + \epsilon D \qquad \text{and} \qquad \delta\phi_n = \epsilon D_{nm}\phi_m$$

Under this transformation $\delta CL = 0$. By Noether's theorem

$$J^\mu = \pi^\mu_n D_{nm}\phi_m = \pi^\mu_2 \phi_1 - \pi^\mu_1 \phi_2 = \phi_1 \partial^\mu \phi_2 - \phi_1 \overleftarrow{\partial}^\mu \phi_2 \equiv \phi_1 \overleftrightarrow{\partial}^\mu \phi_2$$

is a conserved current.

To check that the current is conserved we need equations of motion,

$$\partial^2 \phi_n + 2V' \phi_n = 0$$

Then

$$\partial^\mu J_\mu = \phi_1 \partial^2 \phi_2 - (\partial^2 \phi_1)\phi_2 = \phi_1(2V'\phi_2) - (2V'\phi_1)\phi_2 = 0$$

The conserved charge is

$$Q = \int d^3x \, (\phi_1 \partial_t \phi_2 - \phi_2 \partial_t \phi_1) = \int d^3x \, (\phi_1 \pi_2 - \phi_2 \pi_1) \tag{3.9}$$

For the QFT at a fixed time, say $t = 0$, we have (equal-time) commutation relations

$$i[\pi_n(\vec{x}), \phi_m(\vec{x}\,')] = \delta_{nm}\delta^{(3)}(\vec{a} - \vec{x}\,')$$

and the others (ϕ-ϕ and π-π) vanish. It follows that

$$[Q, \phi_1(\vec{x}\,)] = \int d^3x'\, [\phi_1(\vec{x}\,')\pi_2(\vec{x}\,') - \phi_2(\vec{x}\,')\pi_1(\vec{x}\,'), \phi_1(\vec{x}\,)] = i\phi_2(\vec{x}\,)$$

and similarly

$$[Q, \phi_2(\vec{x}\,)] = -i\phi_1(\vec{x}\,)\,.$$

Together

$$[Q, \phi_n(\vec{x}\,)] = -iD_{nm}\phi_m(\vec{x}\,) = -i\frac{1}{\epsilon}\delta\phi_n(\vec{x}\,) \qquad \text{or} \qquad \delta\phi_n(\vec{x}\,) = i\epsilon[Q, \phi_n(\vec{x}\,)]\,.$$

Since Q is time independent (commutes with H), $\delta\phi_n(\vec{x}\,, t) = i\epsilon[Q, \phi_n(\vec{x}\,, t)]$.

This is in fact a very general result: if $\delta\phi_n = D_{nm}\phi_m$ is a symmetry of $\mathcal{L}$ then the Noether charge Q associated with it has, in the QFT, $\delta\phi_n(\vec{x}\,, t) = i[Q, \phi_n(\vec{x}\,, t)]$.

In order to better understand the physical content of this conserved charge, let's solve the eigensystem for D:

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \chi = \lambda\chi \qquad \Rightarrow \qquad \chi = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm i \end{pmatrix}$$

Therefore we define the field

$$\psi = \frac{\phi_1 - i\phi_2}{\sqrt{2}}$$

We do not define a separate field for $\frac{\phi_1 + i\phi_2}{\sqrt{2}}$ since this is just $\psi^\dagger$ (classically this would be ψ^*, but recall we are dealing with operators in the quantum theory). Then,

$$[Q, \psi] = \tfrac{1}{\sqrt{2}}(i\phi_2 - i(-i\phi_1)) = -\psi \quad \text{and} \quad [Q, \psi^\dagger] = \tfrac{1}{\sqrt{2}}(i\phi 2 + i(-i\phi_1)) = \psi^\dagger$$

So ψ has charge -1 while $\psi^\dagger$ has charge $+1$. This is better seen from the rotation by a finite amount,

$$\begin{aligned} \psi \to \tfrac{1}{\sqrt{2}}\,[(\cos\theta\phi_1 - \sin\theta\phi_2) - i\,(\sin\theta\phi_1 + \cos\theta\phi_2)] \\ = \cos\theta\left(\frac{\phi_1 - i\phi_2}{\sqrt{2}}\right) - i\sin\theta\left(\frac{\phi_1 - i\phi_2}{\sqrt{2}}\right) \end{aligned}$$

or simply

$$\psi \to e^{-i\theta}\psi \qquad \text{and} \qquad \psi^\dagger \to e^{i\theta}\psi^\dagger \tag{3.10}$$

displaying again that $\psi(\psi^\dagger)$ has charge $+1(-1)$. The set of transformations by unitary $n \times n$ matrices form a group, $U(n)$. The transformations in (3.10) are by 1×1 unitary matrices (the phases $e^i\theta$). So the symmetry group is $U(1)$. Since we already knew the symmetry group is $SO(2)$ we see that these groups are really the same (isomorphic). In terms of ψ the classical Lagrangian density is

$$\mathcal{L} = \partial_\mu \psi^* \partial^\mu \psi - V(\psi^* \psi)$$

exhibiting the symmetry under (3.10) quite explicitly.

We can get some further insights by inspecting the action of Q on states. For this we need to expand the fields in terms of creation and annihilation operators but we do not know how to do that for interacting theories (that is, for general "potential" V), nor do we know how to do that for complex fields ψ. So let's take $V = \frac{1}{2}m^2(\phi_1^2 + \phi_2^2)$ and analyze in terms of real fields. We have

$$\phi_n(x) = \int (dk) \left(e^{-ik\cdot x} \alpha_{\vec{k},n} + e^{ik\cdot x} \alpha^\dagger_{\vec{k},n} \right)$$

Plugging this in (3.9) and computing we get

$$Q = i \int (dk) \left(\alpha^\dagger_{\vec{k},2} \alpha_{\vec{k},1} - \alpha^\dagger_{\vec{k},1} \alpha_{\vec{k},2} \right)$$

If we label the particles by an "internal" quantum number, $|\vec{k}, 1\rangle = \alpha^\dagger_{\vec{k},1}|0\rangle$ and $|\vec{k}, 2\rangle = \alpha^\dagger_{\vec{k},2}|0\rangle$, then

$$Q|\vec{k}, 1\rangle = i|\vec{k}, 2\rangle, \quad Q|\vec{k}, 2\rangle = i|\vec{k}, 1\rangle$$

just like the transformation of the fields. This gives relations between probability amplitudes. For this it is important that $[Q, H] = 0$ (which we know holds, but you can verify explicitly for Q and H in terms of creation/annihilation operators). The relation obtained by an infinitesimal transformation is of the form

$$\langle \psi_f | e^{iHt} \left(Q | \psi_i \rangle \right) = \left(\langle \psi_f | Q \right) e^{iHt} | \psi_i \rangle$$

relating the amplitude for $Q|\psi_i\rangle$ to evolve into $|\psi_f\rangle$ in time t, to the amplitude for $|\psi_i\rangle$ to evolve into $Q|\psi_f\rangle$ in the same time. The finite rotation version of this is

$$\langle \psi_f | e^{iHt} | \psi_i \rangle = \langle \psi_f | e^{iHt} e^{-i\theta Q} e^{i\theta Q} | \psi_i \rangle = \langle \psi_f | e^{-i\theta Q} e^{iHt} e^{i\theta Q} | \psi_i \rangle = \langle \psi'_f | e^{iHt} | \psi'_i \rangle$$

where $|\psi'\rangle = e^{i\theta Q}|\psi\rangle$.

Going back to complex fields we have

$$\begin{aligned}\psi(x) &= \int (dk) \left[\left(\frac{\alpha_{\vec{k},1} - i\alpha_{\vec{k},2}}{\sqrt{2}} \right) e^{-ik\cdot x} + \underbrace{\left(\frac{\alpha^\dagger_{\vec{k},1} - i\alpha^\dagger_{\vec{k},2}}{\sqrt{2}} \right)}_{\neq \text{ h.c. of the first term}} e^{ik\cdot x} \right] \\ &= \int (dk) \left(b_{\vec{k}} e^{-ik\cdot x} + c^\dagger_{\vec{k}} e^{ik\cdot x} \right) \\ \psi^\dagger(x) &= \int (dk) \left(c_{\vec{k}} e^{-ik\cdot x} + b^\dagger_{\vec{k}} e^{ik\cdot x} \right)\end{aligned}$$

Notice that

$$[b_{\vec{k}}, b^\dagger_{\vec{k}'}] = [c_{\vec{k}}, c^\dagger_{\vec{k}'}] = 2E_{\vec{k}}(2\pi)^3\delta^{(3)}(\vec{k}' - \vec{k}) \qquad [b_{\vec{k}}, b_{\vec{k}'}] = [c_{\vec{k}}, c_{\vec{k}'}] = 0$$

So these are also creation and annihilation operators, but the one particle states they create are not $|\vec{k}, 1\rangle$ or $|\vec{k}, 2\rangle$, but rather superpositions,

$$|\vec{k}, +\rangle = b^\dagger_{\vec{k}}|0\rangle = \frac{1}{\sqrt{2}}(|\vec{k}, 1\rangle + i|\vec{k}, 2\rangle), \quad \text{and} \quad |\vec{k}, -\rangle = c^\dagger_{\vec{k}}|0\rangle = \frac{1}{\sqrt{2}}(|\vec{k}, 1\rangle - i|\vec{k}, 2\rangle),$$

It is straightforward to get Q in terms of these operators,

$$\alpha_2^\dagger\alpha_1 - \alpha_1^\dagger\alpha_2 = \tfrac{i}{2}(b^\dagger + c^\dagger)(-b + c) - \left(-\tfrac{i}{2}\right)(-b^\dagger + c^\dagger)(b + c) = -i(b^\dagger b - c^\dagger c)$$

so that

$$Q = \int (dk)\,(b^\dagger_{\vec{k}} b_{\vec{k}} - c^\dagger_{\vec{k}} c_{\vec{k}}) = N_+ - N_-$$

where $N_+(N_-)$ counts the number of particles of charge $+(-)$, so that the total charge $Q = N_+ - N_-$ is conserved.

While complex fields can always be recast in terms of pairs of real fields, they can be very useful! So let's discuss briefly the formulation of a field theory directly in terms of ψ and $\psi^\dagger$. Given $\mathcal{L}(\psi, \psi^*, \partial_\mu\psi, \partial_\mu\psi^*)$ (say, as above), how do we obtain the equations of motion? We can do this by varying ϕ_1 and ϕ_2 independently, but how do we make a variation with respect to a complex field? The seemingly dumbest thing to do is to forget that ψ^* is not independent and vary with respect to both ψ and ψ^* (as if tehy were independent). Surprisingly this works. The general argument is this. Suppose you want to find the extremum of a *real* function $F(z, z^*)$,

$$\delta F = f\delta z + f^*\delta z^*\,,$$

for some f. If we naively treat δz and δz^* as independent we obtain the conditions $f = 0 = f^*$. To do this correctly we write $z = x + iy$. For fixed y, $\delta z^* = \delta z$ so that

$f + f^* = 0$; for fixed x, $\delta z^* = -\delta z$ so that $f - f^* = 0$. Combining these conditions we obtain $f = f^* = 0$. This is true for any number, even a continuum, of complex variables. So the equations of motion read

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} - \frac{\partial \mathcal{L}}{\partial \psi} = 0 = \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi^*)} - \frac{\partial \mathcal{L}}{\partial \psi^*}$$

Example:

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

We have

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi^*)} = \partial^\mu \psi, \qquad \frac{\partial \mathcal{L}}{\partial \psi^*} = -m^2 \psi$$

so that

$$(\partial^2 + m^2)\psi = 0\,.$$

This is in accord with the above expansion in terms of plane waves. The real and imaginary parts of ψ satisfy the Klein-Gordon equation. For the Poisson brackets we need the momentum conjugate to ψ,

$$\pi = \frac{\partial \mathcal{L}}{\partial(\partial_t \psi)} = \partial_t \psi^*\,.$$

Then the equal-time commutation relation in the QFT is

$$i[\pi(\vec{x}), \psi(\vec{x}\,')] = \delta^{(3)}(\vec{x} - \vec{x}\,') \quad \Rightarrow \quad i[\partial_t \psi^\dagger(\vec{x}), \psi(\vec{x}\,')] = \delta^{(3)}(\vec{x} - \vec{x}\,')\,.$$

The commutation relation for $\psi^\dagger$ and its conjugate momentum is just the hermitian conjugate of this relation.

Internal symmetries: a non-abelian symmetry example Take now a generalization of the previous example, with N real scalar fields, $\phi_n(x)$, with $n = 1, \ldots, N$, and assume $\mathcal{L}(\phi') = \mathcal{L}(\phi)$ where $\phi'_n = R_{nm}\phi_m$ with R_{nm} real and $R_{nl}R_{ml} = \delta_{nm}$, or, in matrix notation, $\phi' = R\phi$ with $R^T R = RR^T = \mathbb{1}$. For example,

$$\mathcal{L} = \tfrac{1}{2}\partial_\mu \phi_n \partial^\mu \phi_n - V(\phi_n \phi_n) = \tfrac{1}{2}\partial_\mu \phi^T \partial^\mu \phi - V(\phi^T \phi)$$

where in the second step we have rewritten the first expression in matrix form. Note that the set of matrices R that are continuously connected to $\mathbb{1}$ form the group of special orthogonal transformations, $SO(N)$. With $R = \mathbb{1} + \epsilon T$, ϵ infinitesimal, the condition $R^T R = \mathbb{1}$ gives $T^T + T = 0$. That is T is a real antisymmetric and real: there are $\frac{1}{2}N(N-1)$ independent such matrices. For example, for $N = 2$, there is only one such matrix,

$$T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

as in our first example. For $N = 3$ there are three independent matrices which we can take to be

$$T^{12} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad T^{13} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad T^{23} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Here "12" is a label for the matrix T^{12}, etc. We could as well label the matrices T^a with $a = 1, 2, 3$, with $(T^a)_{mn} = \epsilon_{amn}$. In the general case take

$$(T^{ij})_{kl} = \delta^i_k \delta^j_l - \delta^i_l \delta^j_k$$

Then $J^\mu = (\pi^\mu)^T D\phi$ is a set of $\frac{1}{2}N(N-1)$ conserved currents

$$J^{mn}_\mu = \partial_\mu \phi^T T^{mn} \phi = \partial_\mu \phi_k (T^{mn})_{kl} \phi_l = \partial_\mu \phi_m \phi_n - \partial_\mu \phi_n \phi_m = \phi_n \overleftrightarrow{\partial_\mu} \phi_m$$

The matrices T^{mn} don't all commute with each other, so we cannot find simultaneous eigenstates to all of them. We will have more to say about this later and in homework.

Notice the similarity between $J^{\mu mn}$ and $\Delta M^{\mu\nu\lambda}$. In both cases we have the derivative of a field times the field. Notice also the similarity between the numerical coefficients, the matrices $(T^{mn})_{kl}$ and $(\mathcal{I}^{\rho\nu})_{\lambda\sigma}$ encountered in (3.7). This is not an accident. The currents $J^{\mu mn}$ generate $SO(N)$ rotations among the fields ϕ_n while $\Delta M^{\mu\nu\lambda}$ generate $SO(1, 3)$ "rotations" among the fields A^μ.

3.5 Discrete Symmetries

Not every symmetry transformation is continuously connected to the identity. Example,

$$\mathcal{L} = \tfrac{1}{2}(\partial_\mu \phi)^2 - V(\phi^2) \quad \text{is invariant under } \phi(x) \to \phi'(x) = -\phi(x) \tag{3.11}$$

More genrally we can have both transformations that are continuously connected to the identity and transformations that are not. The situation is depicted in Fig. 3.1. The three regions shown form together this example of a group G. One of the disconnected components contains a transformation, K_0 that cannot be reached from $\mathbb{1}$ by a continuous transformation. By definition the connected component of $\mathbb{1}$ contains transformations $U(s)$ that are continuous functions of s and satisfy $U(0) = \mathbb{1}$. We can get to all the elements of G that are in the connected component of K_0 by $U(s)K_0$ (stated without proof). Likewise, the other disconnected component contains a reference element J_0 and all elements are obtained by taking $U((s)J_0$. Since we already understand the physical content of $U(s)$ it

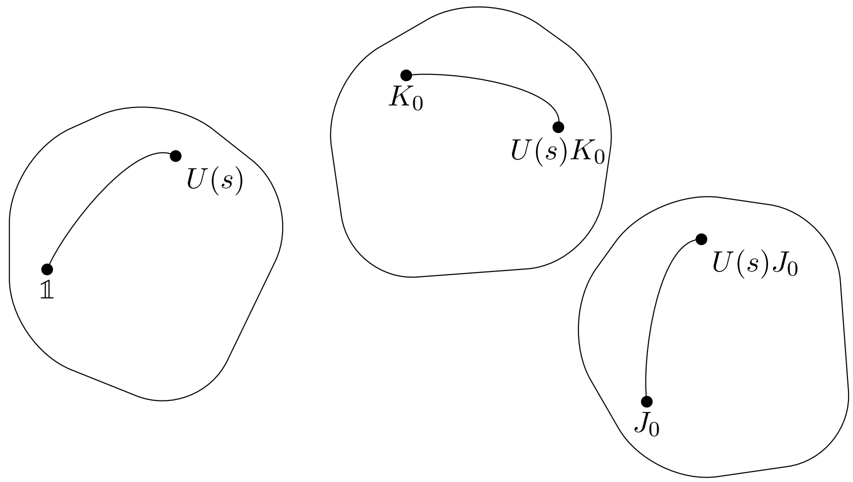


Figure 3.1: Disconnected components of a group of symmetry transformations

suffices to look at a discrete set of symmetry transformations, one per disconnected component, in order to understand this type of group of transformations. These are the *discrete symmetries* we consider now.

Going back to the example in 3.11, we can introduce a unitary operator K such that $K^\dagger\phi(x)K = -\phi(x)$, $K^\dagger K = KK^\dagger = \mathbb{1}$. Note that

$$K^\dagger(K^\dagger\phi(x)K)K = \phi(x) \quad \Rightarrow \quad K^2 = 1 \Rightarrow K^{-1} = K^\dagger = K$$

that is K is hermitian. Strictly speaking we did not show that $K^2 = 1$, since we could as well have $K^2 = e^{i\alpha}$; but we are free to choose the transformation and we make the most convenient choice. Clearly $K^\dagger a_{\vec{k}}K = -a_{\vec{k}}$, $K^\dagger a^\dagger_{\vec{k}}K = -a^\dagger_{\vec{k}}$, so assuming the vacuum is symmetric, $K|0\rangle = |0\rangle$, we have $K|\vec{p}_1,\dots,\vec{p}_n\rangle = (-1)^n|\vec{p}_1,\dots,\vec{p}_n\rangle$. Since n is just the number of particles in the state we can write a representation of K in terms of the number operator, $K = (-1)^N$. Note that we may not have an explicit representation of N in terms of creation and annihilation operators, as in $N = \int(dk)\, a^\dagger_{\vec{k}}a_{\vec{k}}$, because the potential $V(\phi^2)$ generically produces non-linearities ("interactions") in the equation of motion (the case $V(\phi^2) = \frac{1}{2}m^2\phi^2$ is special). Still, it is still true that $K = (-1)^N$, but the number of particles is not conserved. Since $K^\dagger HK = H$, K is conserved, number of particles, N, is conserved mod 2. That means that evolution can change particle number by even numbers; for example, if we call the particle the "chion", χ, then we can have a reaction $\chi+\chi \to \chi+\chi+\chi+\chi$ but not $\chi+\chi \to \chi+\chi+\chi$. Likewise $2\chi \to 84\chi$ and $3\chi \to 11\chi$ are allowed, but not $7\chi \to 16\chi$. The transformations K and $K^2 = \mathbb{1}$

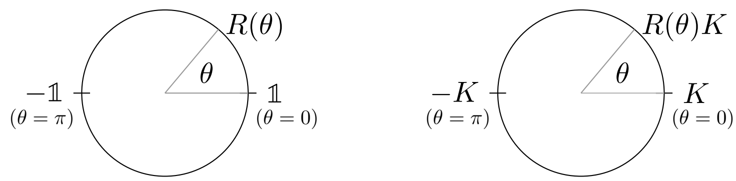


Figure 3.2: Disconnected components of $SO(2) \approx U(1)$

form a group, $G = \{\mathbb{1}, K\}$, isomorphic to $\mathbb{Z}_2$.

3.5.1 Charge Conjugation (C)

Above we saw the example of a Lagrangian with two real fields,

$$\mathcal{L} = \sum_{n=1}^{2} \tfrac{1}{2}\partial^\mu \phi_n \partial_\mu \phi_n - V(\sum_{n=1}^{2}\phi_n^2) = \partial_\mu \psi^* \partial^\mu \psi - V(\psi^*\psi)$$

which is invariant under $\phi_n(x) \to -\phi_n(x)$. But this is not new, it is an $SO(2)$ transformation, a rotation by angle π:

$$\begin{pmatrix}\phi_1(x)\\ \phi_2(x)\end{pmatrix} \to \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\Bigg|_{\theta=\pi}\begin{pmatrix}\phi_1(x)\\ \phi_2(x)\end{pmatrix} = \begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}\phi_1(x)\\ \phi_2(x)\end{pmatrix}$$

In terms of the complex field $\psi \to e^{i\pi}\psi = -\psi$. The more general form of a transformation by a matrix R that preserves the form of the Lagrangian requires only that R be a real orthogonal matrix, that is, that it satisfies $R^T R = RR^T = \mathbb{1}$. This implies $\det(R) = \pm 1$; the matrices $R(\theta)$ with $\det(R(\theta)) = +1$ are the rotations, elements of $SO(2)$. For each of them $R(\theta)K$ with $K = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$ is a matrix with negative determinant, an element of $O(2)$ that is not in $SO(2)$. This is shown in Fig. 3.2, where the left disconnected component is $SO(2)$ and the two components together form $O(2)$.

So let's study K. It is just like above, with $K = (-1)^{N_2}$, and for the linear Klein-Gordon theory $N_2 = \int (dk) a^\dagger_{\vec{k},2} a_{\vec{k},2}$. In terms of the complex field, $K^\dagger(\phi_i - i\phi_2)K = \phi_i + i\phi_2$, so that $\psi \to \psi^\dagger$, that is, K acts as hermitian conjugation.

Since ψ carries charge, $\psi \to \psi^\dagger$ conjugates charge. Hence the name,

$$K^\dagger Q K = -Q \qquad \textit{charge conjugation}$$

Recall also that $b = (a_1 - ia_2)/\sqrt{2}$ and $c = (a_1 + ia_2)/\sqrt{2}$; hence $K^\dagger b_{\vec{k}} K = c_{\vec{k}}$ and $K^\dagger c_{\vec{k}} K = b_{\vec{k}}$. Since $b^\dagger$ ($c^\dagger$) creates particles with charge $+1(-1)$ the action of K is to exchange one particle states of charge $+1$ with one particle states of charge -1. We refer to the $Q=+1$ states as particles and to the $Q=-1$ states as anti-particles, and what we have is that charge conjugation exchanges particles with antiparticles.

In the generic case charge conjugation C has $C|\text{particle}\rangle = |\text{antiparticle}\rangle$ and one still has $C^2=1$ and $C^{-1}=C^\dagger=C$.

3.5.2 Parity (P)

A familiar discrete symmetry in particle mechanics is space inversion,

$$\vec{x} \to -\vec{x}$$

It's QFT version is called *parity*. This is different than the above in that it acts on x^μ. It is part of the Lorentz group, an orthochronous Lorentz transformation (${\Lambda^0}_0 > 0$) with $\det\Lambda = -1$. As we saw earlier, there are four disconnected components of the Lorentz group. And we want representatives from each component:

(i) $\Lambda = \text{diag}(1,1,1,1)$, ${\Lambda^0}_0 > 0, \det\Lambda = +1$ Identity

(ii) $\Lambda = \text{diag}(-1,1,1,1)$, ${\Lambda^0}_0 < 0, \det\Lambda = -1$ Time Reversal (T, $t\to -t$)

(iii) $\Lambda = \text{diag}(1,-1,-1,-1)$, ${\Lambda^0}_0 > 0, \det\Lambda = -1$ Parity (P, $\vec{x}\to-\vec{x}$)

(iv) $\Lambda = \text{diag}(-1,-1,-1,-1)$, ${\Lambda^0}_0 < 0, \det\Lambda = +1$ PT ($x^\mu \to -x^\mu$)

We'll consider T in the next section.

Consider again the Lagrangian in (3.11). Notice that it satisfies $\mathcal{L}(\phi') = \mathcal{L}(\phi)$ with $\phi'(\vec{x},t) = \phi(-\vec{x},t)$. But we could just as well take $\phi'(\vec{x},t) = -\phi(-\vec{x},t)$. If $\mathcal{L}$ has an internal symmetry under $\phi \to U^\dagger\phi U$ and under a parity transformation $\phi \to U_P^\dagger \phi U_P$, then $\widetilde{U}_P = UU_P$ (and also $U_P U$) defines an equally good parity symmetry transformation, $\phi \to \widetilde{U}_P^\dagger \phi \widetilde{U}_P$.

Terminology: if $\phi(\vec{x},t) \to \phi(-\vec{x},t)$ is a symmetry we say that ϕ is a scalar field, as opposed to if $\phi(\vec{x},t) \to -\phi(-\vec{x},t)$ a symmetry, in which case we say it is a *pseudo-scalar* field. If both are symmetries the distinction is immaterial .

The action of parity on states easily understood in terms of creation and annihilation operators (applicable to Klein-Gordon theory, but the result applies more generally):

$$U_P^\dagger \phi(\vec{x},0) U_P = \phi(-\vec{x},0) \quad\Rightarrow\quad \int (dk)\, \left(U_P^\dagger \alpha_{\vec{k}} U_P e^{i\vec{k}\cdot\vec{x}} + \text{h.c.}\right) = \int (dk)\, \left(\alpha_{\vec{k}} e^{-i\vec{k}\cdot\vec{x}} + \text{h.c.}\right)$$

$$\Rightarrow\quad U_P^\dagger \alpha_{\vec{k}}^\dagger U_P = \alpha^\dagger_{-\vec{k}} \quad\Rightarrow\quad U_P|\vec{k}_1,\ldots,\vec{k}_n\rangle = |-\vec{k}_1,\ldots,-\vec{k}_n\rangle$$

More generally, $\mathcal{L}$ is symmetric under parity if it is invariant under a transformation of the form

$$\phi_n(\vec{x},t) \to \phi'_n(\vec{x},t) = R_{nm}\phi_m(-\vec{x},t) \quad n,m = 1,\dots,N$$

for some real matrix R.

Examples:

(i) $\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)^2 - V(\phi^2)$ is P and $\mathbb{Z}_2$ invariant.

(ii) $\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}m^2\phi^2 - g\phi^3$ is P invariant with ϕ a scalar (but not a pseudo-scalar).

(iii) Pions are known to be pseudo-scalars. Here is an example related to $\pi^0 \to \gamma\gamma$. It involves the 4-vector potential A^λ for the electro-magnetic field ($\vec{E} = -\partial_0\vec{A} - \vec{\nabla}A_0$, $\vec{B} = \vec{\nabla}\times\vec{A}$):

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}m^2\phi^2 + g\phi\epsilon^{\mu\nu\lambda\sigma}\partial_\mu A_\nu\partial_\lambda A_\sigma$$

where $\epsilon^{0123} = +1$ is the totally antisymmetric 4-index symbol. You may recall that under parity $\vec{A}(\vec{x},t) \to -\vec{A}(-\vec{x},t)$ and $A_0(\vec{x},t) \to A_0(-\vec{x},t)$. If you do not recall you can either look at Maxwell's equations or you can take a shortcut: from minimal substitution, $\vec{\nabla} \to \vec{\nabla} - e\vec{A}$ one must have $\vec{A}$ transform as $\vec{\nabla}$, and by Lorentz invariance also A_0 as ∂_0. Then the interaction term contributes one of ∂_0 or A_0 and thee of ∂_i or A_i, so $\epsilon^{\mu\nu\lambda\sigma}\partial_\mu A_\nu\partial_\lambda A_\sigma$ is odd under P. Hence $\mathcal{L}$ is P symmetric only if ϕ is a pseudo-scalar, $\phi(\vec{x},t) \to \phi(-\vec{x},t)$. Note that the conclusion is unchanged if A^λ is a *pseudo-vector*, that is, if $\vec{A}(\vec{x},t) \to \vec{A}(-\vec{x},t)$ and $A_0(\vec{x},t) \to -A_0(-\vec{x},t)$ under P.

(iv) Add $h\phi^3$ to the previous example. Now there is no possibility of defining parity that leaves $\mathcal{L}$ invariant.

(v) Consider

$$\mathcal{L} = \sum_{1,2}\left(\tfrac{1}{2}(\partial_\mu\phi_n)^2 - \tfrac{1}{2}m^2\phi_n^2\right) + g\epsilon^{\mu\nu\lambda\sigma}\partial_\mu A_\nu\partial_\lambda A_\sigma(\phi_1^2 - \phi_2^2) + h(\phi_1^3\phi_2 - \phi_1\phi_2^3)\,.$$

This is symmetric under the parity transformation

$$\phi_1(\vec{x},t) \to -\phi_2(-\vec{x},t)$$
$$\phi_2(\vec{x},t) \to \phi_1(-\vec{x},t)$$

In this example $(U_P^\dagger)^2\phi_n(\vec{x},t)U_P^2 = -\phi_n(\vec{x},t)$. It is not true that $U_P^2 = \mathbb{1}$ in general (although for common applications it is).

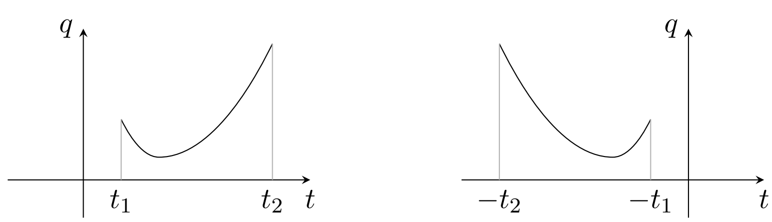
Remark: in the examples above you may feel cheated by the introduction of 4-vector fields, since we have not discussed them in any length. You may instead replace $\epsilon^{\mu\nu\lambda\sigma}\partial_\mu A_\nu \partial_\lambda A_\sigma \to \epsilon^{\mu\nu\lambda\sigma}\partial_\mu\phi_1\partial_\nu\phi_2\partial_\lambda\phi_3\partial_\sigma\phi_4$, for which you need four additional fields. By itself this term in the Lagrangian does not give rise to any dynamics since it is a total derivative. But when multiplied by a function of yet another field it is no longer a total derivative and gives a non-trivial contribution to equations of motion. So consider

$$\mathcal{L} = \sum_{n=1}^{5} \tfrac{1}{2}(\partial_\mu\phi_n)^2 - g\phi_5\epsilon^{\mu\nu\lambda\sigma}\partial_\mu\phi_1\partial_\nu\phi_2\partial_\lambda\phi_3\partial_\sigma\phi_4 \,.$$

Under P, $\phi_n(\vec{x},t) \to (-1)^{\pi_n}\phi_n(-\vec{x},t)$ and the interaction term transforms by a factor of $-(-1)^{\sum_n \pi_n}$.

3.5.3 Time Reversal (T)

In classical mechanics $t \to -t$ is *not* a symmetry. If $L = L(t)$ only though its dependence on $q = q(t)$ and $\dot{q} = \dot{q}(t)$, then while $L(t) \to L(-t)$ is form invariant, so that $q(-t)$ is a solution of equations of motion, the boundary conditions $q(t_1) = q_1$ and $q(t_2) = q_2$ break the symmetry. This simply means that if L has no explicit time dependence and the motion of $q(t)$ from t_1 to t_2 with $q(t_1) = q_1$ and $q(t_2) = q_2$ is allowed, then so is $q(-t)$ from $q(-t_2)$ to $q(-t_1)$:



We'd like to aim for something analogous in QM: an operator that takes a solution of the equation of motion to another,

$$U_T^{-1} q(t) U_T \stackrel{?}{=} q(-t)$$

However, if U_T is a unitary operator we encounter contradictions:

(i) Since we want $U_T^{-1}\dot{q}(t)U_T \stackrel{?}{=} -\dot{q}(-t)$, then $U_T^{-1}p(t)U_T \stackrel{?}{=} -p(-t)$. Then

$$U_T^{-1}[q(t),p(t)]U_T \stackrel{?}{=} \begin{cases} U_T^{-1} i U_T = i & \text{if commutator is computed first} \\ [q(-t), -p(-t)] = -i & \text{if } U_T \text{ applied to operators first} \end{cases}$$

(ii) For any operator, $A(t) = e^{iHt}A(0)e^{-iHt} \equiv e^{iHt}Ae^{-iHt}$ implies

$$A(-t) \stackrel{?}{=} U_T^{-1}A(t)U_T = U_T^{-1}e^{iHt}U_TU_T^{-1}A(0)U_TU_T^{-1}e^{-iHt}U_T = \left(U_T^{-1}e^{iHt}U_T\right)A\left(U_T^{-1}e^{-iHt}U_T\right)$$

but this is also

$$A(-t) = e^{-iHt}Ae^{iHt}\,.$$

Since this is to hold for any A, we must have,

$$U_T^{-1}e^{iHt}U_T \stackrel{?}{=} e^{-iHt}\,.$$

For $t = \epsilon$, infinitesimal, we expand to get $U_T^{-1}(\mathbb{1} + i\epsilon H)U_T \stackrel{?}{=} \mathbb{1} - i\epsilon H$ or $U_T^{-1}HU_T \stackrel{?}{=} -H$. This means the spectrum of H is the same of $-H$. If H is not bounded from above then it is not bounded from below. This does not seem right, that time reflection symmetry requires a negative energy catastrophe!

The solution to these difficulties is to replace an anti-unitary transformation Ω_T for the unitary U_T. Then in (i) and (ii) above $\Omega_T^{-1}i\Omega_T = -i$, and problem solved! While the problem with (ii) above is resolved, it still leads to the requirement

$$\Omega_T^{-1}H\Omega_T = H \qquad\qquad T\text{-invariance}$$

When this is satisfied

$$\begin{aligned}(\psi_f, e^{-iH\Delta t}\psi_i) &= (\psi_f, \Omega_T^{-1}e^{iH\Delta t}\Omega_T\psi_i) \qquad (\Delta t = t_f - t_i)\\ &= (\Omega_T\Omega_T^{-1}e^{iH\Delta t}\Omega_T\psi_i, \Omega_T\psi_f)\\ &= (e^{iH\Delta t}\Omega_T\psi_i, \Omega_T\psi_f)\\ &= (\Omega_T\psi_i, e^{-iH\Delta t}\Omega_T\psi_f)\end{aligned}$$

so the amplitude for ψ_i at $t = t_i$ to evolve to ψ_f at $t = t_f$ is the same as the amplitude for $\Omega_T\psi_f$ at $t = -t_f$ to evolve to $\Omega_T\psi_i$ at $t = -t_i$.

For a single free scalar field,

$$\Omega_T^{-1}\phi(\vec{x}, t)\Omega_T = \eta_T\phi(\vec{x}, -t)\,,$$

where $\eta_T = \pm 1$ (same ambiguity by $\mathbb{Z}_2$ as in case of P). Then, if ϕ satisfies the Klein-Gordon equation we can expand in terms of creation and annihilation operators and

$$\Omega_T^{-1}\alpha_{\vec{p}}\Omega_T = \alpha_{-\vec{p}}, \quad \Omega_T^{-1}\alpha^\dagger_{\vec{p}}\Omega_T = \alpha^\dagger_{-\vec{p}}, \quad \Rightarrow \Omega_T|\vec{k}_1, \ldots, \vec{k}_n\rangle = (\eta_T)^n|-\vec{k}_1, \ldots, -\vec{k}_n\rangle\,.$$

Notice that this is much like P. We can define a PT anti-unitary operator $\Omega_{PT}^{-1}\phi(x^\mu)\Omega_{PT} = \eta_{PT}\phi(-x^\mu)$, which is simpler in that it leaves the states $|\vec{k}_1, \ldots, \vec{k}_n\rangle$ unchanged (save for a factor of $(\eta_{PT})^n = (\eta_P\eta_T)^n$).

3.5.4 CPT Theorem (baby version)

Since Ω_{PT} does not seem to do much, maybe any theory is invariant under PT? Answer: no. Example, a complex scalar field with

$$\mathcal{L} = \partial_\mu \psi^\dagger \partial^\mu \psi - (h\psi^3 + h^*\psi^{\dagger 3} + g\psi^4 + g^*\psi^{\dagger 4}) \,.$$

Then $\Omega_{PT}^{-1}(h\psi^3(x) + g\psi^4(x))\Omega_{PT} = h^*\psi^3(-x) + g^*\psi^4(-x)$ and the interaction part of the Hamiltonian, $H_{\rm int} = \int d^3x\,(h\psi^3 + h^*\psi^{\dagger 3} + g\psi^4 + g^*\psi^{\dagger 4})$ has $\Omega_{PT}^{-1} H_{\rm int}\Omega_{PT} = \int d^3x\,(h^*\psi^3 + h\psi^{\dagger 3} + g^*\psi^4 + g\psi^{\dagger 4})$ which is invariant only if $h^* = h$ and $g^* = g$. (Actually the condition is only $(h^*/h)^4 = (g^*/g)^3$, because one can redefine $\psi \to e^{i\alpha}\psi$ and choose to make g or h real, and when $(h^*/h)^4 = (g^*/g)^3$ both can be made simultaneously real). But note that since $U_C^{-1}\psi(x)U_C = \psi^\dagger(x)$, if we combine C with PT we obtain

$$\Omega_{CPT}^{-1} H \Omega_{CPT} = H \,.$$

For this to work it was crucial that $H^\dagger = H$, as well as that $\mathcal{L}$ is Lorentz invariant. More generally, if we have Lorentz invariance but not hermiticity of the Hamiltonian, we would have $\Omega_{CPT}^{-1} H \Omega_{CPT} = H^\dagger$. For example, if $H_{\rm int} = \int d^3x\,(g\psi^4 + h\psi^{\dagger 4})$ then $\Omega_{CPT}^{-1} H_{\rm int}\Omega_{CPT} = \int d^3x\,(h^*\psi^4 + g^*\psi^{\dagger 4})$.

Notice that we took $\Omega_{PT}^{-1}\psi(x)\Omega_{PT} = \psi(-x)$, which is natural for a complex field: it leaves $b_{\vec{k}}$ and $c_{\vec{k}}$ unchanged, as was the case of $\alpha_{\vec{k}}$ for real fields. The operation that takes $\psi(x)$ into $\psi^\dagger(-x)$ is CPT. If $\Omega_{CPT}^{-1}\phi_n(x)\Omega_{CPT} = \phi_n(-x)$, then

$$\Omega_{CPT}^{-1}\left(\frac{\phi_1(x) - i\phi_2(x)}{\sqrt{2}}\right)\Omega_{CPT} = \frac{\phi_1(-x) + i\phi_2(-x)}{\sqrt{2}} = \psi^\dagger(-x) \,.$$

To obtain a PT transformation $\psi(x) \to \psi(-x)$ from the transformation of the real fields one must have $\Omega_{PT}^{-1}\phi_1(x)\Omega_{PT} = \phi_1(-x)$ and $\Omega_{PT}^{-1}\phi_2(x)\Omega_{PT} = -\phi_2(-x)$. Of course we are free to define anti-unitary operations from the product of a putative Ω_T and any unitary ones, say, U_C or U_P, and investigate then which of these may be a symmetry of the system under consideration. But we want Ω_T to stand for what we physically interpret as time-reversal, which does not involve exchanging anti-particles for particles.

More generally we arrive at the following *CPT theorem*. Consider $\mathcal{L}$ to be a Lorentz invariant function of real scalar fields $\phi_n(x)$ and complex scalar fields $\psi_i(x)$. If $H^\dagger = H$ then $\Omega_{CPT}^{-1} H \Omega_{CPT} = H$. The proof should be obvious by now. Roughly, $H = H(g, \phi_n, \psi_i, \psi_i^\dagger)$; $H^\dagger = H$ implies $H = H(g^*, \phi_n, \psi_i^\dagger, \psi_i)$, and

$$\begin{aligned}
\Omega_{CPT}^{-1} H(g, \phi_n, \psi_i, \psi_i^\dagger)\Omega_{CPT} &= H(\Omega_{CPT}^{-1} g\Omega_{CPT}, \Omega_{CPT}^{-1}\phi_n\Omega_{CPT}, \Omega_{CPT}^{-1}\psi_i\Omega_{CPT}, \Omega_{CPT}^{-1}\psi_i^\dagger\Omega_{CPT}) \\
&= H(g^*, \phi_n, \psi_i^\dagger, \psi_i) \\
&= H(g, \phi_n, \psi_i, \psi_i^\dagger) \,.
\end{aligned}$$

There is an implicit analysis of the monomials that sum up to H, that shows that Lorentz invariance is sufficient to make the change $x^\mu \to -x^\mu$ a formal invariance. Note that for this it is important that the combination PT is a Lorentz transformation with $\det \Lambda = +1$.

The CPT theorem is a surprising consequence of relativistic invariance in consistent (hermitian Hamiltonian) quantum field theory. It implies, for example, that if a theory is CP invariant (which involves unitary transformations) it automatically is T invariant (an anti-unitary transformation). It also gives equality of properties, like mass, of particles and anti-particles. The latter may seem trivial, but it is not once you consider particles that are complex bound states due to srong forces (like the proton, which by CPT has the same mass, magnitude of charge and magnetic moment, as the anti-proton).

Chapter 4

Interaction with External Sources

4.1 Classical Fields and Green's Functions

We start our discussion of sources with classical (non-quantized) fields. Consider

$$\mathcal{L} = \tfrac{1}{2}(\partial_\mu \phi(x))^2 - \tfrac{1}{2}m^2\phi(x)^2 + J(x)\phi(x)$$

Here $J(x)$ is a non-dynamical field, the *source*. The equation of motion for the dynamical field ($\phi(x)$) is

$$(\partial^2 + m^2)\phi(x) = J(x)\,. \tag{4.1}$$

The terminology comes from the more familiar case in electromagnetism. The non-homogeneous Maxwell equations,

$$\vec{\nabla}\cdot\vec{E} = \rho, \qquad \vec{\nabla}\times\vec{B} - \partial_0\vec{E} = \vec{J}$$

have as sources of the electric and magnetic fields the charge and current densities, ρ and $\vec{J}$, respectively. We can go a little further in pushing the analogy. In terms of the electric and vector potentials, A_0 and $\vec{A}$, respectively, the fields are

$$\vec{E} = -\partial_0\vec{A} - \vec{\nabla}A_0, \qquad \vec{B} = \vec{\nabla}\times\vec{A}$$

so that Gauss's law becomes

$$-\partial_0\vec{\nabla}\cdot\vec{A} - \nabla^2 A_0 = \rho$$

In *Lorentz gauge*, $\vec{\nabla}\cdot\vec{A} + \partial_0 A_0 = 0$ this takes the form

$$\partial^2 A_0 = \rho\,.$$

This is exactly for the form of (4.1), for a massless field with the identification A_0 for ϕ and ρ for J. In Lorentz gauge the equations satisfied by the vector potential are again of this form,

$$\partial^2 \vec{A} = \vec{J}\,.$$

The four vector $J^\mu = (\rho, J^i)$ serves as source of the four-vector potential A^μ. Writing

$$\partial^2 A^\mu = J^\mu\,.$$

is reassuringly covariant under Lorentz transformations, as it should be since this is where relativity was discovered! Each of the four components of A^μ satisfies the massless KG equation with source.

Consider turning on and off a localized source. The source is "on," that is non-vanishing, only for $-T < t < T$. Before J is turned on ϕ is evolving as a *free* KG field (meaning free of sources or interactions), so we can think of the initial conditions as giving $\phi(x) = \phi_{\text{in}}(x)$ for $t < -T$, with $\phi_{\text{in}}(x)$ a solution of the free KG equation. We similarly have that for $t > T$ $\phi(x) = \phi_{\text{out}}(x)$ where $\phi_{\text{out}}(x)$ is a solution of the free KG equation. Both $\phi_{\text{in}}(x)$ and $\phi_{\text{out}}(x)$ are solutions of the KG equations for all t but they agree with ϕ only for $t < -T$ and $t > T$, respectively. Our task is to find $\phi_{\text{out}}(x)$ given $\phi_{\text{in}}(x)$ (and of course the source $J(x)$).

Solve (4.1) using Green functions:

$$\phi(x) = \phi_{\text{hom}}(x) + \int d^4y\, G(x-y) J(y) \tag{4.2}$$

where the Green function satisfies

$$(\partial^2 + m^2) G(x) = \delta^{(4)}(x) \tag{4.3}$$

and $\phi_{\text{hom}}(x)$ is a solution of the associated homogeneous equation, which is the KG equation,

$$(\partial^2 + m^2)\phi_{\text{hom}}(x) = 0\,.$$

We can use the freedom in $\phi_{\text{hom}}(x)$ to satisfy boundary conditions. We can determine the Green function by Fourier transform,

$$G(x) = \int \frac{d^4k}{(2\pi)^4}\, e^{ik\cdot x} \widetilde{G}(k) \qquad \delta^{(4)}(x) = \int \frac{d^4k}{(2\pi)^4}\, e^{ik\cdot x}$$

Then

$$(\partial^2 + m^2) \int \frac{d^4k}{(2\pi)^4}\, e^{ik\cdot x} \widetilde{G}(k) = \int \frac{d^4k}{(2\pi)^4}\, e^{ik\cdot x} (-k^2 + m^2) \widetilde{G}(k) = \int \frac{d^4k}{(2\pi)^4}\, e^{ik\cdot x}$$

so that

$$\widetilde{G}(k) = -\frac{1}{k^2 - m^2}\,.$$

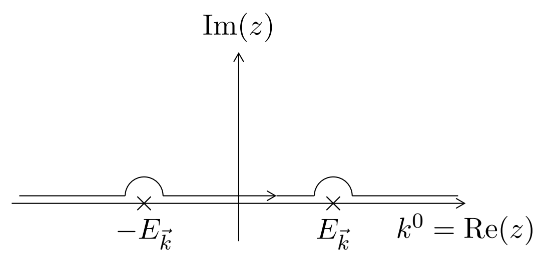
So preliminarily take

$$G(x) = -\int \frac{d^4k}{(2\pi)^4} \frac{e^{ik\cdot x}}{k^2 - m^2}$$

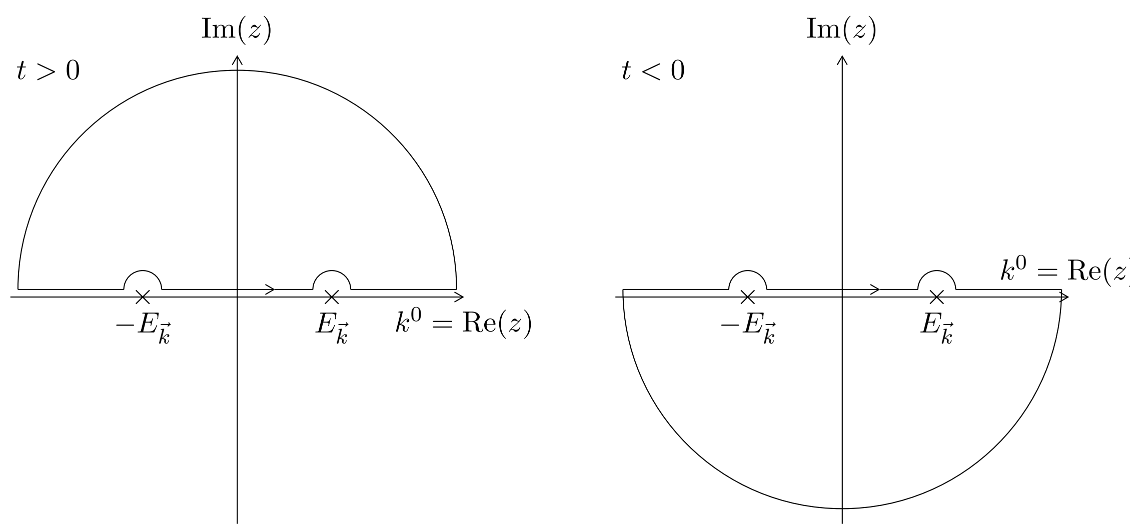
However, note this is ill-defined since the integrand diverges at $k^2 = m^2$. The integral over k^0 diverges at $k^0 = \pm\sqrt{\vec{k}^2 + m^2}$, or $k^0 = \pm E_{\vec{k}}$ for short:

$$\int dk^0 \frac{e^{ik^0 t}}{(k^0 - E_{\vec{k}})(k^0 + E_{\vec{k}})} .$$

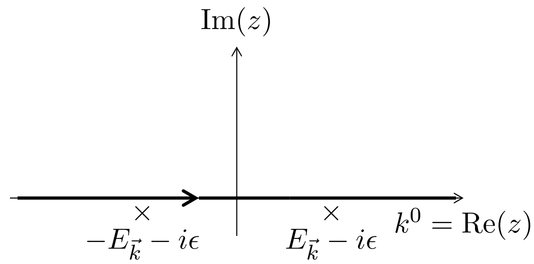
This integral can be thought of as an integral over a complex variable z along a contour on the real axis, $\text{Re}(z) = k^0$, from $-\infty$ to ∞. Then the points $z = \pm E_{\vec{k}}$ are locations of simple poles of the integrand, and we can define the integral by deforming the contour to go either above or below these poles. For example we can take the following contour:



Regardless of which deformation of the contour we choose, the contour can be closed with a semicircle of infinite radius centered at the origin on the upper half-plane if $t > 0$ and in the lower half-plane if $t < 0$:



because he integral along the big semicircle vanishes as the radius of the circle is taken infinitely large. For the choice of contour in this figure no poles are enclosed for $t > 0$ so the integral vanishes. This defined the *advanced* Green's function, $G_{\rm adv}(x) = 0$ for $t > 0$. Alternatively we can "displace" the poles by an infinitesimal amount $-i\epsilon$, with $\epsilon > 0$, so they lie just below the real axis,



Then we have

$$G_{\rm adv}(x) = -\int \frac{d^4k}{(2\pi)^4}\, e^{ik\cdot x} \frac{1}{(k^0 + i\epsilon)^2 - \vec{k}^{\,2} - m^2}$$

Similarly, the *retarded* Green's function is

$$G_{\rm ret}(x) = -\int \frac{d^4k}{(2\pi)^4}\, e^{ik\cdot x} \frac{1}{(k^0 - i\epsilon)^2 - \vec{k}^{\,2} - m^2}$$

with the contour bellow the two poles. It has $G_{\rm ret}(x) = 0$ for $t < 0$. For later use it is convenient to write

$$\begin{aligned} G_{\rm ret}(x) &= -\theta(x^0)\int \frac{d^3k}{(2\pi)^4} e^{-i\vec{k}\cdot\vec{x}} \left[2\pi i \frac{1}{2E_{\vec{k}}}\left(e^{iE_{\vec{k}}t} - e^{-iE_{\vec{k}}t}\right)\right] \\ &= -i\theta(x^0)\int (dk)\,(e^{ik\cdot x} - e^{-ik\cdot x}) \qquad (4.4) \\ G_{\rm adv}(x) &= \theta(-x^0)\int \frac{d^3k}{(2\pi)^4} e^{-i\vec{k}\cdot\vec{x}} \left[2\pi i \frac{1}{2E_{\vec{k}}}\left(e^{iE_{\vec{k}}t} - e^{-iE_{\vec{k}}t}\right)\right] \\ &= i\theta(-x^0)\int (dk)\,(e^{ik\cdot x} - e^{-ik\cdot x}) \qquad (4.5) \end{aligned}$$

We also note that one may choose a contour that goes above one pole and below the other. For example, going below $-E_{\vec{k}}$ and above $+E_{\vec{k}}$ we have

$$\begin{aligned} G_F(x) &= -\int \frac{d^4k}{(2\pi)^4}\, e^{ik\cdot x} \frac{1}{(k^0 - E_{\vec{k}} + i\epsilon)(k^0 + E_{\vec{k}} - i\epsilon)} \\ &= -\int \frac{d^4k}{(2\pi)^4}\, e^{ik\cdot x} \frac{1}{(k^2 - m^2 + i\epsilon)} \qquad (4.6) \\ &= i\left[\theta(x^0)\int (dk)\, e^{-ik\cdot x} + \theta(-x^0)\int (dk)\, e^{ik\cdot x}\right] \end{aligned}$$

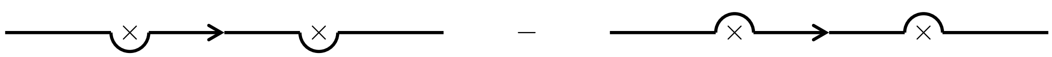
In the solution (4.2), for $x^0 < -T$ the integral over y^0 only has contributions from $x^0 - y^0 < 0$, and for $x^0 > T$ it has contributions only from $x^0 - y^0 > 0$. So we have

$$\begin{aligned}\phi(x) &= \phi_{\rm in}(x) + \int d^4y\, G_{\rm ret}(x-y)J(y)\\ &= \phi_{\rm out}(x) + \int d^4y\, G_{\rm adv}(x-y)J(y)\end{aligned}$$

Hence we can write

$$\begin{aligned}\phi_{\rm out}(x) &= \phi_{\rm in}(x) + \int d^4y\, \left[G_{\rm ret}(x-y) - G_{\rm adv}(x-y)\right]J(y)\\ &= \phi_{\rm in}(x) + \int d^4y\, G^{(-)}(x-y)J(y) \qquad (4.7)\end{aligned}$$

$G^{(-)}$ can be obtained from the difference of (4.4) and (4.5). But more directly we note that



is the same as

The straight segments cancel and we are left with

This gives

$$\begin{aligned}G^{(-)}(x) &= -\int \frac{d^3k}{(2\pi)^4} e^{-i\vec{k}\cdot\vec{x}} \left[2\pi i \frac{1}{2E_{\vec{k}}}\left(e^{iE_{\vec{k}}t} - e^{-iE_{\vec{k}}t}\right)\right]\\ &= -i\int (dk)\left(e^{iE_{\vec{k}}t - i\vec{k}\cdot\vec{x}} - e^{-iE_{\vec{k}}t - i\vec{k}\cdot\vec{x}}\right)\\ &= -i\int \frac{d^4k}{(2\pi)^3}\theta(k^0)\delta(k^2-m^2)\left(e^{ik\cdot x} - e^{-ik\cdot x}\right)\\ &= -i\int \frac{d^4k}{(2\pi)^3}\varepsilon(k^0)\delta(k^2-m^2)e^{ik\cdot x}\end{aligned}$$

where

$$\varepsilon(k^0) = \begin{cases} +1 & k^0 > 0\\ -1 & k^0 < 0\end{cases}$$

You may have noticed that this looks a lot like the function $\Delta_+(x) = [\phi^{(+)}(x), \phi^{(-)}(0)]$. In fact, computing the commutator of free fields at arbitrary times,

$$\begin{aligned}
[\phi(x), \phi(y)] &= \int (dk)(dk')[\alpha_{\vec{k}} e^{-ik\cdot x} + \alpha^\dagger_{\vec{k}} e^{ik\cdot x}, \alpha_{\vec{k}'} e^{-ik'\cdot y} + \alpha^\dagger_{\vec{k}'} e^{ik'\cdot y}] \\
&= \int (dk) \left[e^{ik\cdot(y-x)} - e^{-ik\cdot(y-x)}\right] \\
&= iG^{(-)}(y-x) = -iG^{(-)}(x-y)
\end{aligned}$$

4.2 Quantum Fields

We now consider equation (4.1) for the KG field with a source in the case that the field is an operator on the Hilbert space $\mathcal{F}$. The source $J(x)$ is a a c-number, a *classical source*, and still take it to be localized in space-time. Suppose the system is in some state initially, well before the source is turned on. Let's take the vacuum state for definiteness, although any other state would be just as good. As the system evolves, the source is turned on and then off, and we end up with the system in some final state. Now, in the classical case if we start from nothing and turn the source on and then off radiation is produced, emitted out from the region of the localized source. In the quantum system we therefore expect we will get a final state with any number of particles emitted from the localized source region.

Then $\phi(x) = \phi_{\text{in}}(x)$ is the statement that for $t < -T$ ϕ is the same operator as a solution to the free KG equation. Likewise for $\phi = \phi_{\text{out}}$ for $t > T$. Does this mean $\phi_{\text{out}} = \phi_{\text{in}}$? Surely not, it is not true in the classical case. Since ϕ satisfies the equal time commutation relations, $i[\partial_t\phi(x), \phi(y)] = \delta^{(3)}(\vec{x} - \vec{y})$, etc, so do ϕ_{out} and ϕ_{in}. Since the latter solve the free KG equation, they have expansions in creation and annihilation operators, and the same Hilbert space. More precisely, we can construct a Fock space $\mathcal{F}_{\text{in}}$ out of $\alpha^\dagger_{\text{in}}$, and a space $\mathcal{F}_{\text{out}}$ out of $\alpha^\dagger_{\text{out}}$. These spaces are just Hilbert spaces of the free KG equation, and therefore they are isomorphic, $\mathcal{F}_{\text{in}} \approx \mathcal{F}_{\text{out}} \approx \mathcal{F}_{\text{KG}}$. So there is some linear, invertible operator $S : \mathcal{F}_{\text{KG}} \to \mathcal{F}_{\text{KG}}$, so that $|\psi\rangle_{\text{out}} = S^{-1}|\psi\rangle_{\text{in}}$. Since the states are normalized, S must preserve normalization, so S is unitary, $S^\dagger S = SS^\dagger = 1$. The operator S is called the *S-matrix*.

Let's understand the meaning of $|\psi\rangle_{\text{out}} = S^\dagger|\psi\rangle_{\text{in}}$. The state of the system initially (far past) is $|\psi\rangle_{\text{in}}$. It evolves into a state $|\psi\rangle_{\text{out}} = S^\dagger|\psi\rangle_{\text{in}}$ at late times, well after the source is turned off. We can expand it in a basis of the Fock space, the states $|\vec{k}_1, \ldots, \vec{k}_n\rangle$ for $n = 0, 1, 2, \ldots$ In particular, if the initial state is the vacuum, then the expansion of $S^\dagger|0\rangle$ in the Fock space basis tells us the probability amplitude for emitting any number of particles. More generally

$$_{\text{out}}\langle\chi|\psi\rangle_{\text{in}} = {}_{\text{in}}\langle\chi|S|\psi\rangle_{\text{in}}$$

is the probability amplitude for starting in the state $|\psi\rangle$ and ending in the state $|\chi\rangle$. Note that the left hand side is not to be taken literally as an inner product in the same Hilbert space of free particles (else it would vanish except when the initial and final states are physically identical, *i.e.*, when nothing happens).

Now, $S^\dagger\phi_{\rm in}(x)S$ is an operator acting on out states that satisfies the KG equation. Similarly, if $\pi_{\rm in}(x)=\partial_t\phi_{\rm in}(x)$, $S^\dagger\pi_{\rm in}(x)S$ acts on out states. Moreover, the commutator $[S^\dagger\phi_{\rm in}(x)S, S^\dagger\pi_{\rm in}(y)S = S^\dagger[\phi_{\rm in}(x),\pi_{\rm in}(y)]S = [\phi_{\rm in}(x),\pi_{\rm in}(y)]$ since the commutator is a c-number and $S^\dagger S=1$, and similarly for the other commutators of $S^\dagger\phi_{\rm in}(x)S$ and $S^\dagger\pi_{\rm in}(x)S$. This means that up to a canonical transformation,

$$\phi_{\rm out}(x) = S^\dagger\phi_{\rm in}(x)S\,.$$

Since we are free to choose the out fields in the class of canonical equivalent fields, we take the above relation to be our choice.

There is a simple way to determine S. We will present this now, but the method works only for the case of an external source and cannot be generalized to the case of interacting fields. So after presenting this method we will present a more powerful technique that can be generalized. From Eq. (4.7) we have

$$\begin{aligned} S^\dagger\phi_{\rm in}(x)S &= \phi_{\rm in}(x) + \int d^4y\, G^{(-)}(x-y)J(y) \\ &= \phi_{\rm in}(x) + i\int d^4y\,[\phi_{\rm in}(x),\phi_{\rm in}(y)J(y)]\,. \end{aligned}$$

Recall

$$\begin{aligned} e^ABe^{-A} &= B+[A,B]+\tfrac{1}{2}[A,[A,B]]+\cdots \\ &= B+[A,B] \qquad \text{if } [A,[A,B]]=0. \end{aligned}$$

It follows that

$$\boxed{S=\exp\left[i\int d^4y\,\phi_{\rm in}(y)J(y)\right]} \tag{4.8}$$

It is convenient to normal-order S. Since $[\phi^{(+)}(x),\phi^{(-)}(y)]$ is a c-number we can use

$$e^Ae^B = e^{A+B+\frac{1}{2}[A,B]}$$

which is valid provided $[A,B]$ commutes with both A and B. So using $i\int d^4x\,\phi^{(-)}(x)J(x)$ and $i\int d^4x\,\phi^{(+)}(x)J(x)$ for A and B,

$$S = e^{i\int d^4x\,\phi^{(-)}(x)J(x)}e^{i\int d^4x\,\phi^{(+)}(x)J(x)}e^{\frac{1}{2}\int d^4x\,d^4y\,[\phi^{(-)}(x),\phi^{(+)}(y)]J(x)J(y)}$$

From (2.10),

$$\begin{aligned}\Delta_+(x_2 - x_1) &\equiv [\hat{\phi}^{(+)}(x_1), \hat{\phi}^{(-)}(x_2)] \\ &= \int (dk)\, e^{-ik\cdot(x_2-x_1)} \\ &= \int \frac{d^4k}{(2\pi)^4}\, \theta(k^0)\delta(k^2 - m^2) e^{-ik\cdot(x_2-x_1)}\end{aligned}$$

and introducing the Fourier transform of the source,

$$J(x) = \int \frac{d^4k}{(2\pi)^4}\, e^{ik\cdot x} \widetilde{J}(k)$$

we have

$$\begin{aligned}&\int d^4x\, d^4y\, [\phi_{\text{in}}^{(-)}(x), \phi_{\text{in}}^{(+)}(y)] J(x) J(y) \\ &= -\int \prod_{i=1}^{3} \frac{d^4k_i}{(2\pi)^4} \theta(k_1^0)\delta(k_1^2 - m^2)\widetilde{J}(k_2)\widetilde{J}(k_3) \int d^4x \int d^4y\, e^{ik_2\cdot x} e^{ik_3\cdot y} e^{ik_1\cdot(y-x)} \\ &\qquad = -\int \frac{d^4k}{(2\pi)^4}\, \theta(k^0)\delta(k^2 - m^2)\widetilde{J}(k)\widetilde{J}(-k) = -\int (dk)|\widetilde{J}(k)|^2\,,\end{aligned}$$

where we have used $J^*(x) = J(x) \Rightarrow \widetilde{J}(-k) = \widetilde{J}^*(k)$, and it is understood that $k^0 = E_{\vec{k}}$. Hence,

$$\boxed{S = e^{i\int d^4x\, \phi^{(-)}(x)J(x)} e^{i\int d^4x\, \phi^{(+)}(x)J(x)} e^{-\frac{1}{2}\int (dk)|\widetilde{J}(k)|^2}\,.}$$

As an example of an application, we can compute the probability of finding particles in the final state if we start form no particles initially (emission probability). Start from probability of persistence of the vacuum,

$$|_{\text{out}}\langle 0|0\rangle_{\text{in}}|^2 = |_{\text{in}}\langle 0|S|0\rangle_{\text{in}}|^2 = \exp\left(-\int (dk)|\widetilde{J}(k)|^2\right)\,.$$

Next compute the probability that one particle is produced with momentum $\vec{k}$:

$$|_{\text{out}}\langle \vec{k}\,|0\rangle_{\text{in}}|^2 = |_{\text{in}}\langle \vec{k}\,|S|0\rangle_{\text{in}}|^2 = |_{\text{in}}\langle 0|\alpha_{\vec{k}} S|0\rangle_{\text{in}}|^2\,.$$

where "in" in $\alpha_{\vec{k}}$ is implicit. To proceed we use

$$[\alpha_{\vec{k}}, \phi^{(-)}(x)] = \int (dk')[\alpha_{\vec{k}}, \alpha^\dagger_{\vec{k}'}] e^{ik'\cdot x} = e^{ik\cdot x} \quad (\text{with } k^0 = E_{\vec{k}})$$

so that

$$e^{-i\int\phi^{(-)}J}\alpha_{\vec{k}}e^{i\int\phi^{(-)}J} = \alpha_{\vec{k}} - i\int[\phi^{(-)},\alpha_{\vec{k}}]J = \alpha_{\vec{k}} + i\int d^4x e^{ik\cdot x}J(x) = \alpha_{\vec{k}} + i\widetilde{J}(-k)\,.$$

Hence

$$|_{\text{out}}\langle\vec{k}|0\rangle_{\text{in}}|^2 = |\widetilde{J}(k)|^2\exp\left(-\int(dk)|\widetilde{J}(k)|^2\right)\,.$$

4.2.1 Phase Space

Suppose we want to find the probability of finding a single particle as $t\to\infty$ regardless of its momentum, that is, in any state. We need to sum over all final states consisting of a single particle, once each. Since we have a continuum of states we have to integrate over all $\vec{k}$ with some measure, $\int d\mu(\vec{k})$. Clearly $d\mu(\vec{k})$ must involve d^3k, possibly weighted by a function $f(\vec{k})$. Presumably this function is rotationally invariant, $f = f(|\vec{k}|)$. But we suspect also $d\mu(\vec{k})$ is Lorentz invariant, so $d\mu(\vec{k}) \propto (dk)$, that is, it equals the invariant measure up to a constant. Let's figure this out by counting. Note that how we normalize states matters.

First we determine this via a shortcut, and later we repeat the calculation via a more physical approach (and obtain, of course the same result). We want

$$\sum_n |_{\text{out}}\langle n|0\rangle_{\text{in}}|^2 = \sum_n {}_{\text{in}}\langle 0|n\rangle_{\text{out}}\;{}_{\text{out}}\langle n|0\rangle_{\text{in}} = {}_{\text{out}}\langle 0|\left(\sum_n |n\rangle_{\text{out}}\;{}_{\text{out}}\langle n|_{\text{out}}\right)|0\rangle_{\text{in}}\,,$$

where the sum is restricted over some states. The operator

$$\sum_n |n\rangle_{\text{out}}\;{}_{\text{out}}\langle n|$$

is a projection operator onto the space of "some states," the ones we want to sum in the final state. We have already discussed this projection operator for the case of one particle states when $|\vec{k}\rangle$ is relativistically normalized. It is

$$\int(dk)|\vec{k}\rangle\langle\vec{k}|\,.$$

So we have

$$\text{1-particle emission probability } = \int(dk)|\widetilde{J}(k)|^2\exp\left(-\int(dk)|\widetilde{J}(k)|^2\right)\,.$$

Now we repeat the calculation by counting states. It is easier to count discrete sets of states. We can do so by placing the system in a box of volume $L_xL_yL_z$. Take, say, periodic boundary conditions. Then instead of $\int(dk)(\alpha_{\vec{k}}e^{i\vec{k}\cdot\vec{x}} + \text{h.c.})$, we have a Fourier sum, $\sum_{\vec{k}}(a_{\vec{k}}e^{i\vec{k}\cdot\vec{x}}+\text{h.c.})$. Here $a_{\vec{k}}$ are creation operators with some normalization we will have to sort out. For periodic boundary conditions we must have $k_xL_x = 2\pi n_x$, $k_yL_y = 2\pi n_y$, $k_zL_z = 2\pi n_z$, with n_i integers. We label the one particle states by these, $|\vec{n}\rangle = a^\dagger_{\vec{k}}|0\rangle$, where $\vec{k} = 2\pi\left(\frac{n_x}{L_x}, \frac{n_y}{L_y}, \frac{n_z}{L_z}\right)$. The probability of finding $|\psi\rangle$ in state $|\vec{n}\rangle$ is $|\langle\vec{n}|\psi\rangle|^2$ if both $|\psi\rangle$ and $|\vec{n}\rangle$ are normalized to unity. (Note that this means $[a_{\vec{k}}, a^\dagger_{\vec{k}'}] = \delta_{n_xn'_x}\delta_{n_yn'_y}\delta_{n_zn'_z}$). The probability of finding $|\psi\rangle$ in a 1-particle state is then $\sum_{\vec{n}}|\langle\vec{n}|\psi\rangle|^2$.

Let's be more specific. What is the probability of finding $|\psi\rangle$ in a 1-particle state with momentum in a box $(k_x, k_x+\Delta k_x)\times(k_y, k_y+\Delta k_y)\times(k_z, k_z+\Delta k_z)$? There are $\Delta n_x\Delta n_y\Delta n_z = \frac{L_xL_yL_z}{(2\pi)^3}\Delta k_x\Delta k_y\Delta k_z$ state in the box. For large volume the spacing between values of $\vec{k}$ becomes small, so in the limit of large volume $\Delta k_x \to dk_x$, etc, and the number of particles in the momentum box is

$$\frac{L_xL_yL_z}{(2\pi)^3}\Delta k_x\Delta k_y\Delta k_z \to V\frac{d^3k}{(2\pi)^3}$$

where $V = L_xL_yL_z$ is the volume fo the box. As we switch from discrete to continuum labels for our states we must be careful with their normalization condition. Since on the space of 1-particle states we have

$$\sum_{\vec{n}}|\vec{n}\rangle\langle\vec{n}| = \mathbb{1} \quad\Rightarrow\quad \sum_{\vec{n}}|\vec{n}\rangle\langle\vec{n}|\vec{n}'\rangle = |\vec{n}'\rangle$$

then in the limit, denoting by $|\vec{k}\rangle$ the continuum normalized states,

$$\sum_{\vec{n}}\frac{V\Delta^3k}{(2\pi)^3}|\vec{n}\rangle\langle\vec{n}|\vec{n}'\rangle = |\vec{n}'\rangle \to \int d^3k|\vec{k}\rangle\langle\vec{k}|\vec{k}'\rangle = |\vec{k}'\rangle$$

$$\Rightarrow\quad \frac{V}{(2\pi)^3}\delta_{n_xn'_x}\delta_{n_yn'_y}\delta_{n_zn'_z} \to \delta^{(3)}(\vec{k}-\vec{k}')\,.$$

Putting these elements together, the probability of finding $|\psi\rangle$ in a 1-particle state is

$$\sum_{\vec{n}}|\langle\vec{n}|\psi\rangle|^2 \to \int d^3k|\langle\vec{k}|\psi\rangle|^2\,.$$

Finally, if we want to change the normalization of states so that $\langle\vec{k}'|\vec{k}\rangle = N_{\vec{k}}\delta^{(3)}(\vec{k}-\vec{k}')$ then the probability of finding $|\psi\rangle$ in a 1-particle state is

$$\int\frac{d^3k}{N_{\vec{k}}}|\langle\vec{k}|\psi\rangle|^2\,.$$

For relativistic normalization of states, $N_{\vec{k}} = (2\pi)^3 2E_{\vec{k}}$ and the probability is

$$\int (dk) |\langle \vec{k}|\psi\rangle|^2 \,.$$

This is precisely our earlier result. Denoting the emission probability for n particles in the presence of a localized source $J(x)$ by p_n, we have

$$p_0 = e^{-\xi}, \; p_1 = \xi e^{-xi}, \quad \text{where } \xi = \int (dk) |\widetilde{J}(\vec{k})|^2 \,. \tag{4.9}$$

4.2.2 Poisson

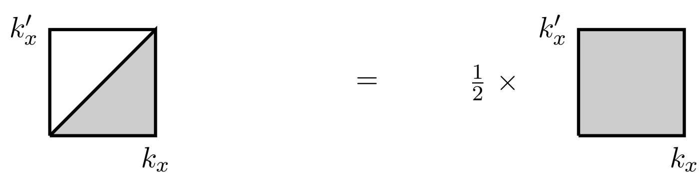
We can now go on to compute p_n for arbitrary n. Start with $n = 2$, that is, emission of two particles.

$$\begin{aligned} e^{-i\int \phi^{(-)}J} \alpha_{\vec{k}} \alpha_{\vec{k}'} e^{i\int \phi^{(-)}J} &= e^{-i\int \phi^{(-)}J} \alpha_{\vec{k}} e^{i\int \phi^{(-)}J} e^{-i\int \phi^{(-)}J} \alpha_{\vec{k}'} e^{i\int \phi^{(-)}J} \\ &= (\alpha_{\vec{k}} + i\widetilde{J}(-k))(\alpha_{\vec{k}'} + i\widetilde{J}(-k')) \end{aligned} \tag{4.10}$$

so that

$$_{\text{out}}\langle \vec{k}\vec{k}'|0\rangle_{\text{in}} = -\widetilde{J}(-k)\widetilde{J}(-k') e^{-\frac{1}{2}\xi} \,.$$

To get the emission probability we must sum over all distinguishable 2-particle states. Since $|\vec{k}\vec{k}'\rangle = |\vec{k}'\vec{k}\rangle$ we must not double count. When we integrate over a box in momentum space, we count twice the state $|\vec{k}\vec{k}'\rangle$ if we sum over k_1 and k_2 with values $k_1 = k$, $k_2 = k'$ and $k_1 = k'$, $k_2 = k$:



Hence

$$p_2 = \tfrac{1}{2}\xi^2 e^{-\xi} \,.$$

The generalization is straightforward:

$$p_n = \frac{1}{n!}\xi^n \exp(-\xi) \,.$$

This is a Poisson distribution! Note that $\sum_n p_n = 1$, that is, there is certainty of finding anything (that is, either no or some particles). The mean of the distribution is $\xi = \int (dk)|\widetilde{J}(\vec{k})|^2$.

4.3 Evolution Operator

We now introduce a more general formalism to derive the same results, but that will be more useful when we consider interacting quantum fields. We want to construct an operator $U(t)$ that gives the connection between the field ϕ and the "in" field $\phi_{\rm in}$:

$$\phi(\vec{x},t) = U^{-1}(t)\phi_{\rm in}(\vec{x},t)U(t)\,. \tag{4.11}$$

Since we are assuming $\phi \to \phi_{\rm in}$ as $t \to -\infty$, we must have

$$U(t) \to 1 \quad \text{as} \quad t \to -\infty\,. \tag{4.12}$$

The S matrix is then

$$S = \lim_{t\to\infty} U(t)\,.$$

The time evolution of ϕ and $\phi_{\rm in}$ are given by

$$\frac{\partial\phi}{\partial t}(\vec{x},t) = i[H(t),\phi(\vec{x},t)] \qquad \frac{\partial\phi_{\rm in}}{\partial t}(\vec{x},t) = i[H_{0\rm in}(t),\phi_{\rm in}(\vec{x},t)] \tag{4.13}$$

Here $H = H_0 + H'$, where H_0 is the *free Hamiltonian* and H' describes interactions. In the present case,

$$H_0 = \int d^3x\,\left[\tfrac{1}{2}\pi^2 + \tfrac{1}{2}(\vec{\nabla}\phi)^2 + \tfrac{1}{2}m^2\phi^2\right] \qquad H'(t) = \int d^3x\, J(\vec{x},t)\phi(\vec{x},t)\,.$$

Also the subscript "in" means the argument has "in" fields. So $H = H(\phi,\pi,J)$ with $H_0 = H_0(\phi,\pi)$ while $H_{0\rm in} = H_{0\rm in}(\phi_{\rm in},\pi_{\rm in})$. From Eq. (4.11) we have

$$\begin{aligned}
&U^{-1}(t)H(\phi(t),\pi(t),J(t))U(t)\\
&\qquad = H(U^{-1}(t)\phi(t)U(t),U^{-1}(t)\pi(t)U(t),U^{-1}(t)J(t)U(t))\\
&\qquad\qquad = H(\phi_{\rm in}(t),\pi_{\rm in}(t),J(t))
\end{aligned} \tag{4.14}$$

and

$$\begin{aligned}
\frac{\partial\phi_{\rm in}}{\partial t}(\vec{x},t) &= \frac{\partial}{\partial t}\left[U\phi(\vec{x},t)U^{-1}(t)\right]\\
&= \frac{dU}{dt}\phi U^{-1} + U\phi\frac{dU^{-1}}{dt} + Ui[H(t),\phi(\vec{x},t)]U^{-1}\\
&= \frac{dU}{dt}U^{-1}\phi_{\rm in} - \phi_{\rm in}\frac{dU}{dt}U^{-1} + i[H_{\rm in}(t),\phi_{\rm in}(\vec{x},t)]\\
&= i[-i\frac{dU}{dt}U^{-1} + H_{\rm in}(t),\phi_{\rm in}(\vec{x},t)]
\end{aligned}$$

Comparing with Eq. (4.13) this is the commutator with $H_{0\text{in}}$ so we must have

$$-i\frac{dU}{dt}U^{-1} + H_{\text{in}}(t) = H_{0\text{in}}(t)$$

or

$$\frac{dU}{dt} = -i(H_{\text{in}} - H_{0\text{in}})U = -iH'_{\text{in}}U \tag{4.15}$$

The solution to this equation with the boundary condition (4.12) gives the S matrix, $S = U(\infty)$. Note that the equation contains only "in" fields, which we know how to handle.

We can solve (4.15) by iteration. Integrating (4.15) form $-\infty$ to t we have

$$U(t) - 1 = -i\int_{-\infty}^{t} dt'\, H'_{\text{in}}(t')U(t')\,.$$

Now use this again, repeatedly,

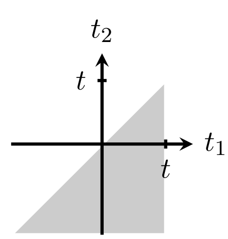
$$\begin{aligned}
U(t) &= 1 - i\int_{-\infty}^{t} dt'\, H'_{\text{in}}(t')\left[1 - i\int_{-\infty}^{t'} dt''\, H'_{\text{in}}(t'')U(t'')\right] \\
&= 1 - i\int_{-\infty}^{t} dt_1\, H'_{\text{in}}(t_1) + (-i)^2\int_{-\infty}^{t} dt_1\int_{-\infty}^{t_1} dt_2 H'_{\text{in}}(t_1)H'_{\text{in}}(t_2)U(t_2) \\
&= \cdots = 1 + \sum_{n=1}^{\infty}(-i)^n\int_{-\infty}^{t} dt_1\int_{-\infty}^{t_1} dt_2\cdots\int_{-\infty}^{t_{n-1}} dt_n\, H'_{\text{in}}(t_1)H'_{\text{in}}(t_2)\cdots H'_{\text{in}}(t_n)\,.
\end{aligned}$$

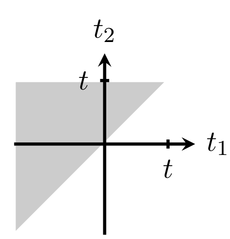
Note that the product $H'_{\text{in}}(t_1)H'_{\text{in}}(t_2)\cdots H'_{\text{in}}(t_n)$ is *time-ordered*, that is, the Hamiltonians appear ordered by $t \geq t_1 \geq t_2 \cdots \geq t_n$. This is the main result of this section.

We can write the result more compactly. Note that

$$\int_{-\infty}^{t} dt_1\int_{-\infty}^{t_1} dt_2 H'_{\text{in}}(t_1)H'_{\text{in}}(t_2) = \int_{-\infty}^{t} dt_2\int_{-\infty}^{t_2} dt_1 H'_{\text{in}}(t_2)H'_{\text{in}}(t_1)$$

The two integrals together cover the region $-\infty < t_1 \leq t$, $-\infty < t_2 \leq t$ of the t_1 *vs* t_2 plane, as can be seen from the followin figures in which the shaded regions correspond to the region of integration of the first and second integrals, respectively:





For any two time dependent operators, $A(t)$ and $B(t)$, we define the *time-ordered product*

$$T(A(t_1)B(t_2)) = \theta(t_1-t_2)A(t_1)B(t_2)+\theta(t_2-t_1)B(t_2)A(t_1) = \begin{cases} A(t_1)B(t_2) & t_1 > t_2 \\ B(t_2)A(t_1) & t_2 > t_1 \end{cases}$$

and similarly when there are more than two operators in the product. Then

$$\int_{-\infty}^{t} dt_1 \int_{-\infty}^{t_1} dt_2 H'_{\rm in}(t_1)H'_{\rm in}(t_2) = \frac{1}{2}\int_{-\infty}^{t} dt_1 \int_{-\infty}^{t} dt_2 T\left(H'_{\rm in}(t_1)H'_{\rm in}(t_2)\right) .$$

For the term with n integrals there are $n!$ orderings of $t_1, \ldots, t_n$ so we obtain

$$U(t) = 1 + \sum_{n=1}^{\infty} \frac{(-i)^n}{n!} \int_{-\infty}^{t} dt \int_{-\infty}^{t} dt_2 \cdots \int_{-\infty}^{t} dt_n\, T(H'_{\rm in}(t_1)H'_{\rm in}(t_2)\cdots H'_{\rm in}(t_n)) . \tag{4.16}$$

or, comparing with $e^z = \sum_{n=0}^{\infty} \frac{1}{n!} z^n$ we write

$$U(t) = T\left[\exp\left(-i\int_{-\infty}^{t} dt'\, H'_{\rm in}(t')\right)\right] .$$

You should keep in mind that the meaning of this expression is the explicit expansion in (4.16). Finally, taking $t \to \infty$,

$$S = T\left[\exp\left(-i\int_{-\infty}^{\infty} dt\, H'_{\rm in}(t)\right)\right] = T\left[\exp\left(-i\int d^4x\, \mathcal{H}'_{\rm in}\right)\right]$$

or, using $\mathcal{L}' = -\mathcal{H}'$,

$$\boxed{S = T\left[\exp\left(i\int d^4x\, \mathcal{L}'_{\rm in}\right)\right] .}$$

Let's use this general result for the specific example we have been discussing, and compare with previous results. Use $\mathcal{L}'_{\rm in} = J(x)\phi_{\rm in}(x)$. Then

$$S = T\left[\exp\left(i\int d^4x\, J(x)\phi_{\rm in}(x)\right)\right] .$$

But we had obtained

$$S = e^{-\frac{1}{2}\xi} :\!\exp\left(i\int d^4x\, J(x)\phi_{\rm in}(x)\right)\!: \tag{4.17}$$

with $\xi = \frac{1}{2}\int (dk)|\widetilde{J}(k)|^2$, as in (4.9). Are these two expression for S the same? To show they are we need some additional machinery, some relation between T-ordered and normal-ordered products.

4.4 Wick's Theorem

In order to answer the question, what is the difference between T-ordered and normal-ordered products, we consider

$$T(\phi(x_1)\phi(x_2)) - {:}\phi(x_1)\phi(x_2){:}$$

Here and below ϕ stands for an "in" field, that is, a free field satisfying the KG equation. For notational conciseness we write ϕ_i for $\phi(x_1$, etc. Letting $\phi = \phi^{(+)} + \phi^{(-)}$ and taking $x_1^0 > x_2^0$ the difference is

$$(\phi_1^{(+)}+\phi_1^{(-)})(\phi_2^{(+)}+\phi_2^{(-)})-(\phi_1^{(+)}\phi_2^{(+)}+\phi_1^{(-)}\phi_2^{(+)}+\phi_2^{(-)}\phi_1^{(+)}+\phi_1^{(-)}+\phi_2^{(-)}) = [\phi_1^{(+)},\phi_2^{(-)}]$$

This is a c-number (equals $\Delta_+(x_2 - x_1)$). Taling the expectation value in the vacuum we obtain, restoring full notation momentarily,

$$T(\phi(x_1)\phi(x_2)) = {:}\phi(x_1)\phi(x_2){:} + \langle 0|T(\phi(x_1)\phi(x_2))|0\rangle\,.$$

Clearly this holds for the more general case of several fields,

$$T(\phi_n(x_1)\phi_m(x_2)) =: {:}\phi_n(x_1)\phi_m(x_2){:} + \langle 0|T(\phi_n(x_1)\phi_m(x_2))|0\rangle\,.$$

Consider next the case of three fields, $T(\phi_1\phi_2\phi_3)$. Without loss of generality take $x_1^0 \geq x_2^0 \geq x_3^0$:

$$T(\phi_1\phi_2\phi_3) = \phi_1\phi_2\phi_3 = \phi_1({:}\phi_2\phi_3{:} + \langle 0|T(\phi_2\phi_3)|0\rangle)$$

Now, $\phi_1 {:}\phi_2\phi_3{:} = (\phi_1^{(+)} + \phi_1^{(-)})\, {:}\phi_2\phi_3{:}$. We need to move $\phi_1^{(+)}$ to the right of all of the $\phi^{(-)}$ operators:

$$\phi_1^{(+)}\, {:}\phi_2\phi_3{:} = {:}\phi_2\phi_3{:}\,\phi_1^{(+)} + [\phi_1^{(+)},\phi_2^{(-)}]\, {:}\phi_3{:} + [\phi_1^{(+)},\phi_3^{(-)}]\, {:}\phi_2{:}$$

where we have used the fact that $[\phi_1^{(+)},\phi_n^{(-)}]$ is a c-number. In fact, it equals $\langle 0|T(\phi_1\phi_n)|0\rangle$. So we have

$$T(\phi_1\phi_2\phi_3) = {:}\phi_1\phi_2\phi_3{:} + {:}\phi_1{:}\langle 0|T(\phi_2\phi_3)|0\rangle + {:}\phi_2{:}\langle 0|T(\phi_1\phi_3)|0\rangle + {:}\phi_3{:}\langle 0|T(\phi_1\phi_2)|0\rangle\,.$$

Of course, ${:}\phi{:} = \phi$, but the notation will more easily generalize below.

More notation, rewrite the above as

$$T(\phi_1\phi_2\phi_3) = {:}\phi_1\phi_2\phi_3{:} + {:}\phi_1\overbrace{\phi_2\phi_3}{:} + {:}\overbrace{\phi_1\phi_2}\phi_3{:} + {:}\overbrace{\phi_1\phi_2\phi_3}{:}\ .$$

where

$$\overbrace{\phi_n\phi_m} = \langle 0|T(\phi_n\phi_m)|0\rangle$$

is called a *contraction*.

Wick's theorem states that

$$
\begin{aligned}
T(\phi_1\cdots\phi_n) = {:}\phi_1\cdots\phi_n{:} &+ \sum_{\substack{\text{pairs}\\(i,j)}} {:}\phi_1\cdots\overline{\phi_i\cdots\phi_j}\cdots\phi_n{:}\\
&+ \sum_{\substack{\text{2-pairs}\\(i,j),(k,l)}} {:}\phi_1\cdots\overline{\phi_i\cdots\phi_j}\cdots\overline{\phi_k\cdots\phi_l}\cdots\phi_n{:}+\cdots\\
&+ \begin{cases} {:}\overline{\phi_1\phi_2}\,\overline{\phi_3\phi_4}\cdots\overline{\phi_{n-1}\phi_n}{:} + \text{all pairings} & n=\text{even}\\ {:}\overline{\phi_1\phi_2}\,\overline{\phi_3\phi_4}\cdots\overline{\phi_{n-2}\phi_{n-1}}\phi_n{:} + \text{all pairings} & n=\text{odd}\end{cases}
\end{aligned} \tag{4.18}
$$

In words, the right hand side is the sum over all possible contractions in the normal ordered product ${:}\phi_1\cdots\phi_n{:}$ (including the term with no contractions). The proof is by induction. We have already demonstrated this for $n=2,3$. Let $W(\phi_1,\ldots,\phi_n)$ stand for the right hand side of (4.18), and assume the theorem is valid for $n-1$ fields. Then assuming $x_1^0 \geq x_2^0 \geq \cdots \geq x_n^0$,

$$
\begin{aligned}
T(\phi_1\cdots\phi_n) &= \phi_1 T(\phi_1\cdots\phi_n)\\
&= \phi_1 W(\phi_2,\ldots,\phi_n)\\
&= (\phi_1^{(+)}+\phi_1^{(-)})W(\phi_2,\ldots,\phi_n)\\
&= \phi_1^{(-)}W(\phi_2,\ldots,\phi_n) + W(\phi_2,\ldots,\phi_n)\phi_1^{(+)} + [\phi_1^{(+)}, W(\phi_2,\ldots,\phi_n)]
\end{aligned}
$$

The first two terms are normal ordered and contain all contractions that do not involve ϕ_1. The last term involves the contractions of ϕ_1 with every field in every term in $W(\phi_2,\ldots,\phi_n)$, therefore all possible contractions. Hence the right hand side contains all possible contractions, which is $W(\phi_1,\ldots,\phi_n)$.

4.4.1 Combinatorics

Let's go back to our computation of S. We'd like to use Wick's theorem to relate $T[\exp(i\int d^4x\,\phi_{\rm in}(x)J(x))]$ to ${:}\exp(i\int d^4x\,\phi_{\rm in}(x)J(x)){:}$, and for this we need a little combinatorics. To make the notation more compact we will continue dropping the "in" label on the "in" fields for the rest of this section. Now, $T[\exp(i\int d^4x\,\phi_{\rm in}(x)J(x))]$ is a sum of terms of the form

$$
\frac{i^n}{n!}\int\prod_{i=1}^n d^4x_i\ J_1\cdots J_n T(\phi_1\cdots\phi_n) = \frac{i^n}{n!}\int\prod_{i=1}^n d^4x_i\ J_1\cdots J_n W(\phi_1,\ldots,\phi_n) \tag{4.19}
$$

Consider the term on the right hand side with one contraction,

$$\frac{1}{n!}\int\prod_{i=1}^{n} d^4x_i \; J_1\cdots J_n \sum_{\substack{\text{pairs}\\(i,j)}} {:}\phi_1\cdots \wick{\c{\phi}_i\cdots\c{\phi}_j}\cdots\phi_n{:}$$
$$= \frac{N_{\text{pairs}}}{n!}\int\prod_{i=1}^{n-2} d^4x_i \; J_1\cdots J_{n-2}\,{:}\phi_1\cdots\phi_{n-2}{:}\int d^4x\,d^4x' J J' \wick{\c{\phi}\c{\phi}'}$$

where N_{pairs} is the number of contractions in the sum, which is the same as the number of pairs (i,j) in the list $1,\ldots,n$. That is, the number of ways of choosing two elements of a list of n objects:

$$N_{\text{pairs}} = \binom{n}{2} = \frac{n!}{2!(n-2)!}$$

Let

$$\zeta = \int d^4x\, d^4y J(x)J(y)\langle 0|T(\phi(x)\phi(y))|0\rangle\,. \tag{4.20}$$

Then

$$1 \text{ pairing terms} = \frac{-\zeta}{2}\frac{i^{n-2}}{(n-2)!}\int\prod_{i=1}^{n-2} d^4x_i \; J_1\cdots J_{n-2}\,{:}\phi_1\cdots\phi_{n-2}{:}\,.$$

We want to repeat this calculation for the rest of the terms on the right hand side of (4.19). For this we need to count the number of terms for a given number of contractions. Now, k contractions involve $2k$ fields. Now, there are $\binom{n}{2k}$ ways of picking $2k$ fields out of n, and we need to determine how many distinct contractions one can make among them. We can figure this out by inspecting a few simple cases. For $k=1$

$$\wick{\c1 1 \c1 2} \qquad \to \text{ one contraction}$$

and for $k=2$

$$\wick{\c1 1 \c1 2 \c2 3 \c2 4} \quad \wick{\c1 1 \c2 2 \c1 3 \c2 4} \quad \wick{\c1 1 \c2 2 \c2 3 \c1 4} \qquad \to 3 \text{ contractions.}$$

Instead of continuing in this explicit way, we analyze the $k=3$ case using inductive reasoning:

$$\wick{\c1 1 \c1 2}\times(k=2) \quad \wick{\c1 1 \c1 3}\times(k=2) \quad \cdots \quad \wick{\c1 1 \c1 6}\times(k=2) \qquad \to 5\times 3 \text{ contractions.}$$

By induction the arbitrary k case has $(2k-1)!!$ pairings:

$$\Big[(2k-1)\text{-pairings:}\wick{\c1 1 \c1 j}\Big]\times\Big[(k-1)\text{-case: } (2k-3)!!\Big] = (2k-1)!!$$

So the terms with k contractions give

$$\underbrace{\frac{1}{n!}\frac{n!}{(n-2k)!(2k)!}(2k-1)!!}_{\frac{1}{(n-2k)!}\frac{1}{2^k k!}}\, i^{n-2k}\int\prod_{i=1}^{n-2k} d^4x_i J_1\cdots J_{n-2k} :\phi_1\cdots\phi_{n-2k}: \underbrace{\left(i^2\int d^4x\, d^4x'\, JJ'\phi\phi'\right)^k}_{(-1)^k\zeta^k}$$

$$= \frac{(-\zeta/2)^k}{k!}\;\frac{i^m}{m!}\int\prod_{i=1}^{m} d^4x_i\, J_1\cdots J_m :\phi_1\cdots\phi_m: \qquad \text{with } m = n-2k.$$

Aha! We recognize this as a term in an exponential expansion. Considering all the terms in he expansion of $T[\exp(i\int d^4x\,\phi_{\rm in}(x)J(x))]$, the coefficient of $\frac{i^m}{m!}\int\prod_{i=1}^m d^4x_i\,J_1\cdots J_m :\phi_1\cdots\phi_m:$ (fixed m) is $\sum_{k=0}^\infty \frac{1}{k!}(-\zeta/2)^k = \exp(-\frac{1}{2}\zeta)$, and is independent of m, so it factors out. We are left with

$$e^{-\zeta/2}\sum_{m=0}^\infty \frac{i^m}{m!}\int\prod_{i=1}^m d^4x_i\,J_1\cdots J_m :\phi_1\cdots\phi_m: = e^{-\zeta/2} :\exp\left(i\int d^4x\,J(x)\phi(x)\right):$$

That is

$$S = T\left[\exp\left(i\int d^4x\,\phi_{\rm in}(x)J(x)\right)\right] = e^{-\zeta/2} :\exp\left(i\int d^4x\,J(x)\phi(x)\right):$$

This will equal our previous expression for S in (4.17) if $\zeta = \xi$. To answer this we need to know more about the T-ordered product in the definition of ζ in (4.20).

4.5 Scalar Field Propagator

Let

$$\Delta_F(x,y) \equiv \langle 0|T(\phi(x)\phi(y))|0\rangle$$

where $\phi(x)$ is a real, scalar, free field satisfying the Klein-Gordon equation. This is the quantity we need for the computation fo ζ, but it is also important for several other reasons, so we spend some time investigating it.

First of all, $i\Delta_F(x,y)$ is a Green's function for the Klein-Gordon equation; see (4.3). To verify this claim compute directly. Taking ∂_μ to be the derivative with respect to x^μ keeping y^μ fixed and using $(\partial^2+m^2)\phi(x) = 0$ we have

$$\begin{aligned}(\partial^2+m^2)\Delta_F(x,y) &= (\partial^2+m^2)\left[\theta(x^0-y^0)\langle 0|\phi(x)\phi(y)|0\rangle + \theta(y^0-x^0)\langle 0|\phi(y)\phi(x)|0\rangle\right]\\ &= \frac{\partial^2}{\partial x^{02}}\theta(x^0-y^0)\langle 0|\phi(x)\phi(y)|0\rangle + \frac{\partial^2}{\partial x^{02}}\theta(y^0-x^0)\langle 0|\phi(y)\phi(x)|0\rangle\\ &\quad + 2\frac{\partial}{\partial x^0}\theta(x^0-y^0)\langle 0|\frac{\partial\phi(x)}{\partial x^0}\phi(y)|0\rangle + 2\frac{\partial}{\partial x^0}\theta(y^0-x^0)\langle 0|\phi(y)\frac{\partial\phi(x)}{\partial x^0}|0\rangle\end{aligned}$$

Now using $d\theta(x)/dx = \delta(x)$ and $\partial_0\phi(x) = \pi(x)$ the last line above is

$$2\delta(x^0 - y^0)\langle 0|[\pi(x), \phi(y)]|0\rangle = -2i\delta^{(4)}(x - y)$$

The line above that gives

$$\frac{\partial}{\partial x^0}\delta(x^0 - y^0)\langle 0|[\phi(x), \phi(y)]|0\rangle = -\delta(x^0 - y^0)\langle 0|[\pi(x), \phi(y)]|0\rangle = i\delta^{(4)}(x - y)$$

Combining these we have

$$(\partial^2 + m^2)\Delta_F(x, y) = -i\delta^{(4)}(x - y)$$

As we saw in Sec. 4.1 Green functions are not unique, since one can always add solutions to the homogenous Klein-Gordon equation to obtain a new Green's function. We must have

$$i\Delta_F(x, y) = -i\int \frac{d^4k}{(2\pi)^4} e^{ik\cdot(x-y)} \frac{1}{k^2 - m^2}$$

with some prescription for the contour of integration. Note that $\Delta_F(x, y) = \Delta_F(x-y)$ depends only on the difference $x - y$, which is as expected from homogeneity pof space-time.

Recall that $[\phi^{(+)}(x), \phi^{(-)}(y)] = \langle 0|T\phi(x)\phi(y)|0\rangle$ for $x^0 > y^0$. More generally

$$\langle 0|T\phi(x)\phi(y)|0\rangle = \theta(x^0 - y^0)[\phi^{(+)}(x), \phi^{(-)}(y)] + \theta(y^0 - x^0)[\phi^{(+)}(y), \phi^{(-)}(x)]$$

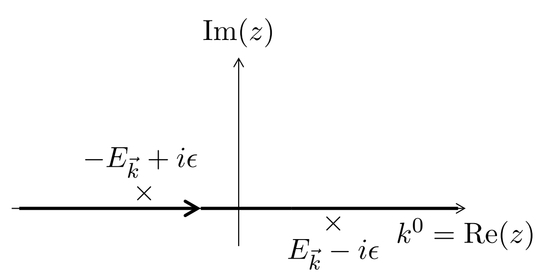
Recall also that

$$[\phi^{(+)}(x), \phi^{(-)}(y)] = \int (dk) e^{-iE_{\vec{k}}(x^0-y^0)+i\vec{k}\cdot(\vec{x}-\vec{y})}$$

Without loss of generality and to simplify notation we set $y^\mu = 0$. We have

$$\langle 0|T\phi(x)\phi(0)|0\rangle = \theta(x^0)\int (dk) e^{-iE_{\vec{k}}x^0+i\vec{k}\cdot\vec{x}} + \theta(-x^0)\int (dk) e^{iE_{\vec{k}}x^0-i\vec{k}\cdot\vec{x}}$$

Comparing with Eq. (4.6) we see this is precisely $-iG_F(x)$, which was obtained by taking a contour that goes below $-E_{\vec{k}}$ and above $E_{\vec{k}}$,



So we have

$$\boxed{\langle 0|T\phi(x)\phi(y)|0\rangle = \int \frac{d^4k}{(2\pi)^4}\, e^{-ik\cdot(x-y)}\frac{i}{k^2-m^2+i\epsilon}} \tag{4.21}$$

This will be of much use later. We will refer to this as the *two-point function* of the KG field, and the Fourier transform, $1/(k^2-m^2+i\epsilon)$, as the KG *propagator*.

We can finally return to the question of relating ζ to ξ:

$$\begin{aligned}
\zeta &= \int d^4x\, d^4y\, J(x)J(y)\, \langle 0|T\phi(x)\phi(y)|0\rangle \\
&= \int d^4x\, d^4y \int \frac{d^4k_1}{(2\pi)^4}\frac{d^4k_2}{(2\pi)^4} e^{ik_1\cdot x} e^{ik_2\cdot y}\widetilde{J}(k_1)\widetilde{J}(k_2) \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot(x-y)}\frac{i}{p^2-m^2+i\epsilon} \\
&= \int \frac{d^4p}{(2\pi)^4}\widetilde{J}(p)\widetilde{J}(-p)\frac{i}{p^2-m^2+i\epsilon}
\end{aligned}$$

Perform the integral over p^0, assuming $\widetilde{J}(p)\widetilde{J}(-p) = |\widetilde{J}(p)|^2$ vanishes as $|p^0| \to \infty$, which is justified by our assumption that the source is localized in space-time. Closing the contour on the upper half of the complex p^0 plane we pick the pole at $-E_{\vec{k}}$ so that

$$\zeta = 2\pi i \int \frac{d^3p}{(2\pi)^4}|\widetilde{J}(p)|^2\frac{i}{-2E_{\vec{k}}} = \int (dk)|\widetilde{J}(p)|^2 = \xi\,.$$

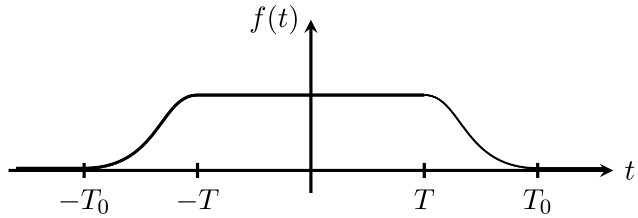
Chapter 5

Elementary Processes

We want to extend the previous discussion to the case where fields interact among themselves, rather than with an external source. The aim is to give an expression for the S-matrix in terms of "in" fields, as before. We will see that uit is convenient to express the matrix elements of S in terms of Green's functions, that is, vacuum expectation values of time-ordered products of elementary fields, as in $\langle 0|T\phi(x_1)\cdots\phi(x_n)|0\rangle$.

As before the S matrix connects $|\text{in}\rangle$ states to $|\text{out}\rangle$ states, $|\text{out}\rangle = S^{-1}|\text{in}\rangle$, and correspondingly ϕ_{in} fields to ϕ_{out} fields. But we can no longer say that $\phi(x) \to \phi_{\text{in}}(x)$ as $t \to -\infty$ (nor $\phi(x) \to \phi_{\text{out}}(x)$ as $t \to +\infty$) with $\phi_{\text{in,out}}(x)$ free fields (a *free field* is one without interactions, e.g., it satisfies the KG equation).

Before we explain that in more detail let us better understand the role of $|\text{in}\rangle$ states (and ϕ_{in} fields). In a collision process we start with particles that are wildly separated, so interactions between them can be initially neglected. As particles approach each other the interactions can no longer be neglected, they "turn on." So one could think of the situation by replacing H' (the interaction part of the Hamiltonian) by $f(t)H'$, where $f(t)$ is a smooth function that turns on slowly (adiabatically), then stays on for some long period over which the collision takes place, say, $f(t) = 1$ for $-T < t < T$, and then turns off slowly again, $f(t) = 0$ $|t| > T_0$ and $f(t)$ smoothly decreasing (increasing) in $T < t < T_0$ ($-T_0 < t < -T$):



We want both T and T_0 to be arbitrarily large, and we want $T_0 \gg T$ in the process to avoid sudden changes that can introduce extraneous effects, e.g., pair production.

We will use this model, but it is not quite general enough. The reason is that if the interaction is turned off we may not be able to describe $|\text{in}\rangle$ and $|\text{out}\rangle$ states of interest, namely, bound states that arise because of the interaction. For example, we may want to study collisions of an electron with an H atom due to electromagnetic interactions. But it is the electromagnetic interactions that binds a p and an e into an H atom. More poignant is the case of collisions of protons by the strong interactions when it is these interactions that keeps quarks bound in protons. The idea of collision theory is that oen can set up a muck theory of free particles that happen to have the same mass (and other quantum numbers, e.g., spin) as the bound states. These are the $|\text{in}\rangle$ and $|\text{out}\rangle$ states. For very early (or late) times these states describe the evolution of the particles that later (earlier) participate in the collision. And the S-matrix uses information in the interacting theory to connect the pre- and post-collision states. For theories without bound states we can use the simpler approximation of turning on and off the interaction via $f(t)H'$. Remarkably the expression for the S-matrix obtained via this simplified treatment is the same as in a more complete and rigorous analysis that does not employ it.

If we adiabatically turn on and off the interactions then we can use our previous approach:

$$\phi(x) = \phi_{\text{in}}(x) + \int d^4y\, G_{\text{ret}}(x-y)J(y) \tag{5.1}$$

But now

$$(\partial^2 + m^2)\phi = \frac{\mathcal{L}'}{\partial\phi}$$

(e.g., if $\mathcal{L}' = g\phi^3 + \lambda\phi^4$ then $\frac{\mathcal{L}'}{\partial\phi} = 3g\phi^2 + 4\lambda\phi^4$). So for $J(x)$ use $f(t)\frac{\mathcal{L}'}{\partial\phi}$. Note that in the absence of $f(t)$ this would not work, the "source" would not be localized in time. Now, we had two ways of obtaining the S-matrix from this. One was

$$S^\dagger\phi_{\text{in}}(x)S = \phi_{\text{in}}(x) + \int d^4y\, [\phi_{\text{in}}(x), \phi_{\text{in}}(y)]J(y)\,.$$

But now $J(y)$ depends on $\phi(x)$ so it does not commute with $\phi_{\text{in}}(x)$. This makes it harder to solve for S using this method.

The second method used $\phi = U^\dagger(t)\phi_{\text{in}}(x)U(t)$ and constructed $U(t)$ in terms of $\mathcal{H}'$. This will work. The result was, and still is,

$$S = T\left[\exp\left(i\int d^4x\, \mathcal{L}'_{\text{in}}\right)\right]$$

which we can use, but now with, say, $\mathcal{L}'_{\text{in}} = g\phi_{\text{in}}^3 + \lambda\phi_{\text{in}}^4$.

However, the above discussion has to be modified, as we will see shortly, because in general one cannot take $\phi(x) \to \phi_{\text{in}}(x)$ as $t \to -\infty$.

5.1 Källen-Lehmann Spectral Representation

Here we will see that we cannot take $\phi(x) \to \phi_{\rm in}(x)$ as $t \to -\infty$. We study $[\phi(x), \phi(y)]$ and compare with $[\phi_{\rm in}(x), \phi_{\rm in}(y)]$. In particular,

$$\langle 0|[\phi(x), \phi(y)]|0\rangle = \sum\!\!\!\!\!\!\int_n \left(\langle 0|\phi(x)|n\rangle\langle n|\phi(y)|0\rangle - x \leftrightarrow y\right)$$

Now

$$\langle 0|\phi(x)|n\rangle = \langle 0|e^{i\hat{P}\cdot x}\phi(0)e^{-i\hat{P}\cdot x}|n\rangle = e^{-ip_n\cdot x}\langle 0|\phi(0)|n\rangle$$

where $\hat{P}^\mu|n\rangle = p_n^\mu|n\rangle$, $\hat{P}^\mu|0\rangle = 0$, and

$$\begin{aligned}
\langle 0|[\phi(x), \phi(y)]|0\rangle &= \sum\!\!\!\!\!\!\int_n \left(e^{-ip_n\cdot(x-y)}\langle 0|\phi(0)|n\rangle\langle n|\phi(0)|0\rangle - x \leftrightarrow y\right) \\
&= \sum\!\!\!\!\!\!\int_n \int d^4k\, \delta^{(4)}(p_n - k)\left(e^{-ik\cdot(x-y)} - e^{ik\cdot(x-y)}\right)|\langle 0|\phi(0)|n\rangle|^2 \\
&= \int \frac{d^4k}{(2\pi)^3}\left(e^{-ik\cdot(x-y)} - e^{ik\cdot(x-y)}\right)\rho(k)
\end{aligned}$$

where in going from the first to the second line we introduce a factor of $1 = \int d^4k\, \delta^{(4)}(p_n - k)$ and in the last line we defined

$$\rho(k) \equiv \sum\!\!\!\!\!\!\int_n (2\pi)^3\delta^{(4)}(p_n - k)|\langle 0|\phi(0)|n\rangle|^2 = \sigma(k^2)\theta(k^0)\,.$$

The last equality follows from (i) Lorentz invariance and (ii) $p_n^0 > 0$.

Now compare this with the case of free fields. We have computed this, but it is easy to derive from above: $|n\rangle$ is only the one particle states $\vec{p}$, $\sum\!\!\!\!\!\int_n$ is $\int(dp)$, and

$$\langle 0|\phi_{\rm in}(0)|\vec{p}\rangle = \int(dk)\langle 0|\alpha_{\vec{k}}\alpha^\dagger_{\vec{p}}|0\rangle = 1\,.$$

Then

$$\rho(k) = \int(dp)(2\pi)^3\delta^{(4)}(p-k) = \int d^4p\,\theta(p^0)\delta(p^2-m^2)\delta^{(4)}(p-k) = \theta(k^0)\delta(k^2-m^2)\,,$$

and

$$\begin{aligned}
\langle 0|[\phi_{\rm in}(x), \phi_{\rm in}(y)]|0\rangle &= \int \frac{d^4k}{(2\pi)^3}\,\theta(k^0)\delta(k^2 - m^2)\left(e^{-ik\cdot(x-y)} - e^{ik\cdot(x-y)}\right) \\
&= \int \frac{d^4k}{(2\pi)^3}\,\varepsilon(k^0)\delta(k^2 - m^2)e^{-ik\cdot(x-y)} \equiv i\Delta(x - y; m)
\end{aligned}$$

Hence

$$\begin{aligned}\langle 0|[\phi(x),\phi(y)]|0\rangle &= \int \frac{d^4k}{(2\pi)^3}\,\sigma(k^2)\varepsilon(k^0)e^{-ik\cdot(x-y)}\\ &= \int \frac{d^4k}{(2\pi)^3}\int_0^\infty d\bar{m}^2\,\delta(k^2-\bar{m}^2)\sigma(k^2)\varepsilon(k^0)e^{-ik\cdot(x-y)}\\ &= \int_0^\infty d\bar{m}^2\sigma(\bar{m}^2)\int \frac{d^4k}{(2\pi)^3}\varepsilon(k^0)\delta(k^2-\bar{m}^2)e^{-ik\cdot(x-y)}\end{aligned}$$

or

$$\boxed{\langle 0|[\phi(x),\phi(y)]|0\rangle = i\int_0^\infty d\bar{m}^2\sigma(m^2)\Delta(x-y;\bar{m})} \tag{5.2}$$

This is the *Källen-Lehmann representation.*

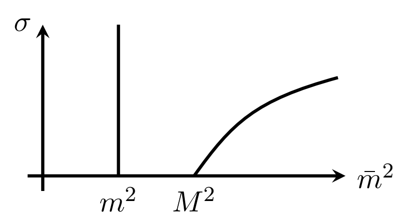
Now we separate the contributions to σ of 1-particle states from $\geq$ 2-particles states. We will assume that $p^0 \geq M > m$ for $\geq$ 2-particle states. For two free particles $p^0 \geq 2m$. If interacting we expect $p^0 \geq 2m-\varepsilon$ where ε is some interaction energy; we are assuming $\varepsilon < m$. If we had $\varepsilon > m$ then the 2-particle energy would be smaller than m and the 1-particle "state" is not a state because it can decay into a lower energy state. Then, if in fact we could demand $\phi(x)\to\phi_{\rm in}(x)$ as $t\to-\infty$ we should have

$$\langle 0|\phi(0)|\vec{p}\rangle \stackrel{?}{=} \langle 0|\phi_{\rm in}(0)|\vec{p}\rangle = 1\,,$$

and therefore

$$\sigma(\bar{m}^2) \stackrel{?}{=} \delta(m^2-\bar{m}^2)+\sigma(\bar{m}^2)\theta(\bar{m}-M)\,. \tag{5.3}$$

This is illustrated in the following figure:



Inserting (5.3) in (5.2) gives

$$\langle 0|[\phi(x),\phi(y)]|0\rangle \stackrel{?}{=} i\Delta(x-y;m)+i\int_{M^2}^\infty d\bar{m}^2\sigma(m^2)\Delta(x-y;\bar{m})\,.$$

Taking $\partial/\partial x^0$ of this, and then the limit $y^0\to x^0$ we obtain on the left hand side the equal time commutator $[\pi(x),\phi(y)] = -i\delta^{(3)}(\vec{x}-\vec{y})$. Then note that on the right hand we can do this again since $i\Delta(x-y;m) = \langle 0|[\phi_{\rm in}(x),\phi_{\rm in}(y)]|0\rangle$. This gives

$$1 \stackrel{?}{=} 1+\int_{M^2}^\infty d\bar{m}^2\,\sigma(\bar{m}^2)$$

If this equation holds then $\sigma(\bar{m}^2) = 0$ for $m > M$. That is $\sigma(\bar{m}^2) = \delta(\bar{m}^2 - m^2)$, which means $\langle 0|\phi(0)|n\rangle = 0$ for any state $|n\rangle$ which has ≥ 2 particles. This then gives $\phi(x) = \phi_{\text{in}}(x)$ for all times which makes $\phi(x)$ a free field. We conclude that we cannot demand $\phi(x) \to \phi_{\text{in}}(x)$ as $t \to -\infty$. Assume instead

$$\phi(x) \to Z^{\frac{1}{2}}\phi_{\text{in}}(x) \qquad \text{as } t \to -\infty$$

Then, repeating the steps above,

$$1 = Z + \int_{M^2}^{\infty} d\bar{m}^2\, \sigma(\bar{m}^2)$$

so that $\sigma > 0$ requires $0 \leq Z < 1$ (that $Z \geq 0$ is from it being $(Z^{\frac{1}{2}})^2$). Similarly, we assume $\phi(x) \to Z^{\frac{1}{2}}\phi_{\text{out}}(x) \qquad \text{as } t \to +\infty$.

We conclude the section with some useful observations. Uniqueness of the vacuum state gives $|0\rangle_{\text{out}} = |0\rangle_{\text{in}} = |0\rangle$. (In principle one can have a relative phase, $|0\rangle_{\text{out}} = e^{i\alpha}|0\rangle_{\text{in}}$ but we conventionally set $\alpha = 0$). Since we are assuming the 1-particle states are stable, they are eigenstates of the Hamiltonian, so they evolve simply, by a phase, e^{-iEt}. Hence $|\vec{k}\rangle_{\text{out}} = |\vec{k}\rangle_{\text{in}}$ (up to a constant phase that we conventionally set to zero). Now $\langle 0|\phi(x)|\vec{k}\rangle = \langle 0|\phi(0)|\vec{k}\rangle e^{-ik\cdot x}$ so that the prescription to evaluate at $t \to -\infty$ in order to compare with the corresponding expectation value of $\phi_{\text{in}}(x)$ is superfluous, and similarly for $t \to \infty$ and expectation values of $\phi_{\text{out}}(x)$. So we have

$$\langle 0|\phi(x)|\vec{k}\rangle = Z^{\frac{1}{2}}\langle 0|\phi_{\text{in}}(x)|\vec{k}\rangle = Z^{\frac{1}{2}}\langle 0|\phi_{\text{out}}(x)|\vec{k}\rangle\,.$$

We collect some basic results for the S-matrix:

$$\phi_{\text{in}}(x) = S\phi_{\text{out}}(x)S^{-1}\,, \qquad |\psi\rangle_{\text{in}} = S|\psi\rangle_{\text{out}}\,,$$
$${}_{\text{out}}\langle\chi|\psi\rangle_{\text{in}} = {}_{\text{out}}\langle\chi|S|\psi\rangle_{\text{out}} = {}_{\text{in}}\langle\chi|S|\psi\rangle_{\text{in}}\,.$$

For ψ, χ the vacuum or 1-particle states

$$\langle 0|S|0\rangle = \langle 0|0\rangle = 1\,,$$
$$\langle\vec{k}|S|\vec{k}'\rangle = \langle\vec{k}|\vec{k}'\rangle = (2\pi)^3 2E_{\vec{k}}\,\delta^{(3)}(\vec{k}' - \vec{k})\,.$$

Finally, if $U = U(a^\mu, \Lambda)$ is a Poincare transformation, covariance means

$$USU^{-1} = S\,.$$

It is often stated in textbooks that $\phi(x) \to Z^{\frac{1}{2}}\phi_{\text{in}}(x)$ cannot hold in the strong sense. That is, that it can only hold for separate matrix elements. Else we'd have, the argument goes, for example, $[\phi(x),\phi(y)] \to Z[\phi_{\rm in}(x),\phi_{\rm in}(y)]$ for non-equal, early times, and since for "in" fields this is a c-number, one would be able to argue that $\phi(x)$ is a free field. I think this is overkill. Obviously $\phi(x0$ and $\phi_{\rm in}$ are different, one is an interacting field and one is not. One can produce multiple particle states out of the vacuum —that's the statement that $\sigma > 0$— the other cannot. The statement that $\phi(x) \to Z^{\frac{1}{2}}\phi_{\rm in}(x)$ at $t \to -\infty$ is useful because it gives us the correct way of relating wildly separated initial state (single)-particles created by ϕ to those created by $\phi_{\rm in}$. To make sense of this we need particles that are truly separated, which means we have to consider wave-packets rather than plane waves. We will comment on this when we discuss the S-matrix for multi-particle states.

5.2 LSZ reduction formula: stated

LSZ stands for Lehmann, Symanzik and Zimmermann. The LSZ formula gives the probability amplitude for scattering any number of particles into any number of particles:

$$
{}_{\rm out}\langle \vec p_1,\ldots,\vec p_l|\vec k_1,\ldots,\vec k_n\rangle_{\rm in} = (iZ^{-\frac{1}{2}})^{n+l}\int\prod_{i=1}^{l} d^4y_i \int\prod_{j=1}^{n} d^4x_j\, e^{i\sum_{i=1}^{l}p_i\cdot y_i - i\sum_{j=1}^{n}k_j\cdot x_j}
$$
$$
\times\prod_{i=1}^{l}(\partial_{y_i}^2+m^2)\prod_{j=1}^{n}(\partial_{x_j}^2+m^2)\langle 0|T\left(\phi(y_1)\cdots\phi(y_l)\phi(x_1)\cdots\phi(x_n)\right)|0\rangle\,. \quad (5.4)
$$

Comments:

(i) Computing S matrix elements reduced to computing Green's functions,

$$
G^{(n)}(x_1,\ldots,x_n) = \langle 0|T\left(\phi(x_1)\cdots\phi(x_n)\right)|0\rangle\,.
$$

(ii) One can do a more general treatment in term of 1-particle wave-packets. Since the LSZ formula is multilinear in the plane waves for the in and out states, the result amounts to replacing $e^{-ik_i\cdot x_i)} \to f_i(x_i)$ and $e^{ip_i\cdot y_i)} \to f_i^*(y_i)$.

(iii) Integrating by parts $(\partial_x^2+m^2)e^{\pm ip\cdot x)} = -(p^2-m^2)e^{\pm ip\cdot x)} = 0$. The result has to be interpreted with care. Let

$$
\int\prod_{i=1}^{n} d^4x_i\, e^{-i\sum_{i=1}^{n}k_i\cdot x_i}G^{(n)}(x_1,\ldots,x_n) = (2\pi)^4\delta^{(4)}(\sum_i k_i)\widetilde G^{(n)}(k_1,\ldots,k_n)\,. \quad (5.5)
$$

That we always have a $\delta^{(4)}(\sum_i k_i)$ follows form translation invariance. We can change variables to the differences $x_{i+1}-x_i$ together with the center of mass $X = \sum_i x_i$. Then since $G^{(n)}$ does not depend on X we will have $\int d^4X e^{-iR\cdot\sum k_i}$ times the rest. Now, $\widetilde{G}^{(n)}(k_1,\ldots,k_n)$ is defined for arbitrary four vectors, $k_1,\ldots,k_n$, not necessarily satisfying the *on-shell* condition $k_i^2 = m^2$; we say that k_i is *off-shell* if $k_i^2 \neq m^2$, or alternatively, that the "energies," k_i^0, are arbitrary, not given by $\pm E_{\vec{k}_i}$. Incidentally, the on/off-shell language is simply short for the momentum being on/off the *mass-shell*. It may be that $\widetilde{G}^{(n)}(k_1,\ldots,k_n)$ has simple poles as $k^0 \to \pm E_{\vec{k}}$. In fact, by Lorentz invariance the poles must be paired, appearing as poles in $k^2 - m^2$. These poles cancel the zeroes from $\prod(\partial^2 + m^2)$ and the S-matrix element is just the residue:

$$
{}_{\text{out}}\langle \vec{p}_1,\ldots,\vec{p}_l|\vec{k}_1,\ldots,\vec{k}_n\rangle_{\text{in}} = (iZ^{-\frac{1}{2}})^{n+l}\int\prod_{i=1}^{l} d^4y_i \int\prod_{j=1}^{n} d^4x_j\, e^{i\sum_{i=1}^{l}p_i\cdot y_i - i\sum_{j=1}^{n}k_j\cdot x_j}
$$

$$
\times\int\prod_{k=1}^{l+n} d^4q_k e^{i\sum_{i=1}^{l}q_i\cdot y_i + i\sum_{j=1}^{n}q_j\cdot x_j}\prod_{i=1}^{l+n}(-q_k^2+m^2)(2\pi)^4\delta^{(4)}(\sum_i k_i)\widetilde{G}^{(n)}(q_1,\ldots,q_{n+l})
$$

$$
= (-iZ^{-\frac{1}{2}})^{n+l}(2\pi)^4\delta^{(4)}(\sum_j k_j - \sum_i p_i)
$$

$$
\lim_{\substack{p_i^2\to m^2\\ k_j^2\to m^2}}\prod_{i=1}^{l}(p_i^2 - m^2)\prod_{j=1}^{n}(k_j^2 - m^2)\widetilde{G}^{(n)}(k_1,\ldots,k_n,-p_1,\ldots,-p_l)\,.
$$

and note that each factor of $p^2 - m^2$ comes with a $1/(iZ^{\frac{1}{2}})$.

(iv) It is therefore useful to summarize this in terms of a *scattering amplitude*,

$$
\begin{aligned}
i\mathcal{A} &= i\mathcal{A}(k_1,\ldots,k_n;p_1,\ldots,p_l)\\
&= \lim_{\substack{p_i^2\to m^2\\ k_j^2\to m^2}}\prod_{i=1}^{l}\frac{(p_i^2 - m^2)}{iZ^{\frac{1}{2}}}\prod_{j=1}^{n}\frac{(k_j^2 - m^2)}{iZ^{\frac{1}{2}}}\widetilde{G}^{(n)}(k_1,\ldots,k_n,-p_1,\ldots,-p_l)
\end{aligned}
$$

so that

$$
{}_{\text{out}}\langle \vec{p}_1,\ldots,\vec{p}_l|\vec{k}_1,\ldots,\vec{k}_n\rangle_{\text{in}} = (2\pi)^4\delta^{(4)}(\sum_j k_j - \sum_i p_i)\, i\mathcal{A}\,.
$$

5.3 S-matrix: perturbation theory

Continuing to present results without justification (which will be given later) we now give the Green's functions (or *correlators* or *n-point functions* in terms of "in" fields:

$$\begin{aligned} G^{(n)}(x_1,\dots,x_n) &= \langle 0|T(\phi(x_1)\cdots\phi(x_n))|0\rangle \\ &= \langle 0|T(\phi_{\rm in}(x_1)\cdots\phi_{\rm in}(x_n)e^{-i\int d^4x\mathcal{H}'_{\rm in}})|0\rangle \end{aligned} \tag{5.6}$$

Expanding the exponential, writing $\epsilon\mathcal{H}'_{\rm in}$ for $\mathcal{H}'_{\rm in}$, with $\epsilon=1$, just a counting device, and retaining up to some power in ϵ, say ϵ^N, we are approximating $G^{(n)}$ as a perturbative expansion. If $\mathcal{H}'_{\rm in}$ is written explicitly in terms of some parameters, and these can be considered as small, we are then approximating $G^{(n)}$ as an expansion in powers of these small parameters. For example, with $\mathcal{L}=\mathcal{L}_0+\mathcal{L}'$, $\mathcal{L}_0=\frac{1}{2}(\partial_\mu\phi)^2-\frac{1}{2}m^2\phi^2$ and $\mathcal{L}'=-\mathcal{H}'=-(\lambda/4!)\phi^4$, then we are expanding $G^{(n)}$ in powers of λ, the *coupling constant*.

Let's see this explicitly in this example. Compute $G^{(4)}(x_1,\dots,x_4)$:

0th order We take the 1 in the expansion fo the exponential:

$$G_0^{(4)}(x_1,\dots,x_4)=\langle 0|T(\phi_{\rm in}(x_1)\cdots\phi_{\rm in}(x_4))|0\rangle$$

To make the notation more compact we drop the label "in" for now and use ϕ_1 for $\phi(x_1$, etc . Using Wick's theorem,

$$\begin{aligned} G_0^{(4)}(x_1,\dots,x_4) &= \langle 0|T(\phi_1\cdots\phi_4)|0\rangle \\ &= \langle 0|\,:\phi_1\cdots\phi_4: + \contraction{:}{\phi}{_1}{\phi}:\phi_1\phi_2\phi_3\phi_4: + \contraction{:}{\phi}{_1\phi_2}{\phi}:\phi_1\phi_2\phi_3\phi_4: + \contraction{:}{\phi}{_1\phi_2\phi_3}{\phi}:\phi_1\phi_2\phi_3\phi_4: \\ &\quad + \contraction{:\phi_1}{\phi}{_2}{\phi}:\phi_1\phi_2\phi_3\phi_4: + \contraction{:\phi_1}{\phi}{_2\phi_3}{\phi}:\phi_1\phi_2\phi_3\phi_4: + \contraction{:\phi_1\phi_2}{\phi}{_3}{\phi}:\phi_1\phi_2\phi_3\phi_4: \\ &\quad + \contraction{:}{\phi}{_1}{\phi}\contraction{:\phi_1\phi_2}{\phi}{_3}{\phi}:\phi_1\phi_2\phi_3\phi_4: + \contraction{:}{\phi}{_1\phi_2}{\phi}\contraction[2ex]{:\phi_1}{\phi}{_2\phi_3}{\phi}:\phi_1\phi_2\phi_3\phi_4: + \contraction[2ex]{:}{\phi}{_1\phi_2\phi_3}{\phi}\contraction{:\phi_1}{\phi}{_2}{\phi}:\phi_1\phi_2\phi_3\phi_4:\,|0\rangle \end{aligned}$$

Only the last line is non-vanishing, the previous two have normal ordered operators acting on the vacuum. The last line gives

$$G_0^{(4)}(x_1,\dots,x_4)=\contraction{}{\phi}{_1}{\phi}\contraction{\phi_1\phi_2}{\phi}{_3}{\phi}\phi_1\phi_2\phi_3\phi_4+\contraction{}{\phi}{_1\phi_2}{\phi}\contraction[2ex]{\phi_1}{\phi}{_2\phi_3}{\phi}\phi_1\phi_2\phi_3\phi_4+\contraction[2ex]{}{\phi}{_1\phi_2\phi_3}{\phi}\contraction{\phi_1}{\phi}{_2}{\phi}\phi_1\phi_2\phi_3\phi_4$$

Now, we compute $\widetilde{G}^{(4)}$ and then $i\mathcal{A}$:

$$\begin{aligned} \int\prod_{n=1}^{4}d^4x_n\,e^{-i\sum k_n\cdot x_n}G_0^{(4)}(x_1,\dots,x_4) &= \int d^4x_1\,d^4x_2 e^{-ik_1\cdot x_1-ik_2\cdot x_2}\contraction{}{\phi}{_1}{\phi}\phi_1\phi_2 \\ &\quad\times\int d^4x_3\,d^4x_4 e^{-ik_3\cdot x_3-ik_4\cdot x_4}\contraction{}{\phi}{_3}{\phi}\phi_3\phi_4+\text{permutations.} \end{aligned}$$

We need

$$
\begin{aligned}
\int d^4x_1\, d^4x_2 e^{-ik_1\cdot x_1 - ik_2\cdot x_2} \int \frac{d^4q}{(2\pi)^4} e^{-iq\cdot(x_1-x_2)} \frac{i}{q^2-m^2+i\epsilon}& \\
= \int \frac{d^4q}{(2\pi)^4}\frac{i}{q^2-m^2+i\epsilon}\int d^4x_1\, e^{-ix_1\cdot(k_1+q)}\int d^4x_2\, e^{-ix_2\cdot(k_2-q)}& \\
= \int \frac{d^4q}{(2\pi)^4}\frac{i}{q^2-m^2+i\epsilon}(2\pi)^4\delta^{(4)}(k_1+q)\,(2\pi)^4\delta^{(4)}(k_2-q)& \\
= (2\pi)^4\delta^{(4)}(k_1+k_2)\frac{i}{k_1^2-m^2}&
\end{aligned}
$$

There is no $i\epsilon$ in the last step since we have performed the integral over q^0. Using this above we have

$$
\begin{aligned}
\widetilde{G}^{(4)}(k_1,\ldots,k_4) &= \frac{i}{k_1^2-m^2}\frac{i}{k_3^2-m^2}(2\pi)^4\delta^{(4)}(k_1+k_2)(2\pi)^4\delta^{(4)}(k_3+k_4) + \text{perms} \\
&= (2\pi)^4\delta^{(4)}(k_1+k_2+k_3+k_4)\left[\frac{i}{k_1^2-m^2}\frac{i}{k_3^2-m^2}(2\pi)^4\delta^{(4)}(k_1+k_2) + \text{perms}\right]
\end{aligned}
$$

Moreover, the scattering amplitude is

$$
\begin{aligned}
&i\mathcal{A}(k_1,k_2;p_1,p_2) = \\
&\lim_{\substack{p_i^2\to m^2\\ k_j^2\to m^2}}\left(\frac{p_1^2-m^2}{i}\right)\left(\frac{p_2^2-m^2}{i}\right)\left(\frac{k_1^2-m^2}{i}\right)\left(\frac{k_2^2-m^2}{i}\right)\widetilde{G}^{(4)}(k_1,\ldots,k_4) = 0\,,
\end{aligned}
$$

which makes sense since at this order in the expansion there is no interaction so there is no scattering.

1st order We now expand the exponential to linear order so that

$$
G_1^{(4)}(x_1,\ldots,x_4) = \langle 0|T(\phi_1\cdots\phi_4(-i)\int d^4x\,\tfrac{\lambda}{4!}\phi^4(x))|0\rangle
$$

To calculate this we use Wick's theorem. We know from experience gained above that we need the terms with all fields contracted. Let's distinguish terms like

$$
\overline{\phi_1\phi_2}\,\overline{\phi_3\phi_4}\,(-i)\int d^4x\,\frac{\lambda}{4!}\overline{\phi(x)\phi(x)}\,\overline{\phi(x)\phi(x)} \quad\text{or}\quad \overline{\phi_1\phi_2}\,(-i)\int d^4x\frac{\lambda}{4!}\phi_3\phi_4\phi(x)\phi(x)
$$

where at least two of the $\phi_1,\ldots,\phi_4$ are contracted among themselves, from terms like

$$
-i\frac{\lambda}{4!}\int d^4x\phi_1\phi_2\phi_3\phi_4\phi(x)\phi(x)\phi(x)\phi(x) \tag{5.7}
$$

We call the first kind *disconnected*, the second *connected*. The reason for the terminology will become clear when we introduce a graphical representation of these contractons. We say the fields $\phi_1, \ldots, \phi_4$ are *external* while the ones that appear from inserting powers of the Hamiltonian into the time ordered product are *internal*.

For disconnected terms each contraction $\overbracket{\phi_i \phi_j}$ of a pair of external fields will give a single factor of $(2\pi)^4\delta^{(4)}(k_i + k_j) i(k_i^2 - m^2)^{-1}$, and the rest of the factors in that term will be independent of k_i and k_j. Then, when computing the amplitude $i\mathcal{A}$ we'll have

$$\lim_{k_{i,j}^2 \to m^2} (k_i^2 - m^2)(k_j^2 - m^2) \frac{i}{k_i^2 - m^2} (2\pi)^4 \delta^{(4)}(k_i + k_j) \times (k_{i,j}\text{-independent}) = 0$$

To get a non-vanishing amplitude we look in connected terms. Note that the one in (5.7) is but one of 4! contractions of this type, and they all give the same result. So we have

$$\begin{aligned} G^{(4)}_{1,\text{conn}}(x_1, \ldots, x_4) &= -i\lambda \int d^4x\, \overbracket{\phi(x_1)\phi(x)}\overbracket{\phi(x_2)\phi(x)}\overbracket{\phi(x_3)\phi(x)}\overbracket{\phi(x_4)\phi(x)} \\ &= -i\lambda \int d^4x \prod_{n=1}^{4} \left(\int \frac{d^4q_n}{(2\pi)^4} e^{-iq_n\cdot(x_n - x)} \frac{i}{q_n^2 - m^2 + i\epsilon} \right) \\ &= -i\lambda \int \prod_{n=1}^{4} \frac{d^4q_n}{(2\pi)^4} e^{-iq_n\cdot(x_n - x)} (2\pi)^4 \delta^{(4)}(\sum_{n=1}^{4} q_n) \prod_{n=1}^{4} \frac{i}{q_n^2 - m^2 + i\epsilon} \end{aligned}$$

from which we read off

$$\widetilde{G}^{(4)}_{1,\text{conn}}(k_1, \ldots, k_4) = -i\lambda \prod_{n=1}^{4} \frac{i}{q_n^2 - m^2 + i\epsilon}$$

We now use the LSZ formula to compute the scattering amplitude. This entails removing the four propagators and multiplying by Z^{-2}. However, Z itself has a perturbative expansion, with $Z = 1 + \mathcal{O}(\lambda)$, so to the order we are working we can set $Z = 1$. We obtain

$$i\mathcal{A}(k_1, k_2; p_1, p_2) = -i\lambda\,, \tag{5.8}$$

that is

$$_{\text{out}}\langle \vec{p}_1 \vec{p}_2 | \vec{k}_1 \vec{k}_2 \rangle_{\text{in}} = -i\lambda (2\pi)^4 \delta^{(4)}(k_1 + k_2 - p_1 - p_2)$$

5.3.1 Graphical Representation

A graphical representation gives an effective way of communicating and organizing these calculations. Consider the Green's function that we would need to compute the scattering amplitude at p-th order in λ in the perturbative expansion

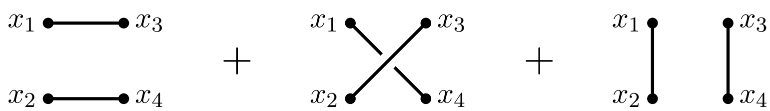
$$(-i\lambda)^p \langle 0|T\left(\phi(x_1)\cdots\phi(x_4)\,\frac{\phi^4(y_1)}{4!}\cdots\frac{\phi^4(y_p)}{4!}\right)|0\rangle \tag{5.9}$$

In using Wick's theorem to compute this is expanded into a sum of many terms. Each term in the sum is represented by a diagram, and the set of all diagrams, is constructed by drawing:

1. An endpoint of a line for each $x_1,\ldots,x_4$. We call these lines *external*.
2. A point, or *vertex*, from which four lines originate for each $y_1,\ldots,y_p$.
3. Connections between the loose ends of the lines, so that all points $x_1,\ldots,y_p$ are connected with lines (with one line emerging from the x's and four from the y's).

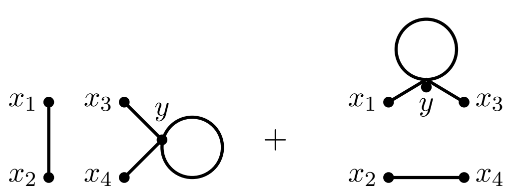
We associate with each line a factor of $\overbracket{\phi(z_a)\phi(z_b)}$, where z_a and z_b are the two points connected by the line. To each vertex we associate a $-i\Lambda$. Finally there is a combinatorial factor arising from equivalent contractions, to compensate for less than 4! possible equivalent contractions; see example and fuller explanation below.

Let's recover the results of our 0-th and 1st order calculations. At lowest order, $p=0$, so there are not vertices, only four endpoints of external lines:



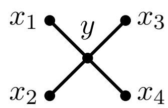
$$= \overbracket{\phi_1\phi_3}\,\overbracket{\phi_2\phi_4} \quad + \quad \overbracket{\phi_1\phi_4}\,\overbracket{\phi_2\phi_3} \quad + \quad \overbracket{\phi_1\phi_2}\,\overbracket{\phi_3\phi_4}$$

These are disconnected terms, and the diagrams are *disconnected diagrams*. As such they give $\mathcal{A}=0$. Now at first order in perturbation theory, the $p=1$ term, we have disconnected diagrams,



$$+\ \cdots = -i\lambda\tfrac{1}{2}\,\overbracket{\phi_1\phi_y}\,\overbracket{\phi_3\phi_y}\,\overbracket{\phi_y\phi_y}\,\overbracket{\phi_2\phi_4} + \cdots$$

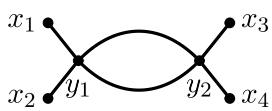
and one connected diagram,



$$= -i\lambda\, \overline{\phi_1\phi_y}\; \overline{\phi_2\phi_y}\; \overline{\phi_3\phi_y}\; \overline{\phi_4\phi_y}$$

Note that the first disconnected diagram has a combinatorial factor of $\frac{1}{2}$, as indicated. This can be seen as follows. Starting from (5.9) with $p = 1$, and insisting that ϕ_1 is contracted with ϕ_2 and both ϕ_3 and ϕ_4 are contracted with ϕ_y's, we see that there is only one possible contraction of ϕ_1 with ϕ_2, then we have 4 ways of contracting ϕ_3 with ϕ_y^4, which ϕ_y^3 un-contracted, and finally we have 3 ways of contracting ϕ_4 with ϕ_y^3. The last step leaves ϕ_y^2 which can give a single contraction of ϕ_y with ϕ_y. That is, there are 4×3 contractions. This times the pre-factor of $\frac{1}{4!}$ gives the symmetry factor of $\frac{1}{2}$.

Here is another example of a combinatorial factor, now for a connected diagram, from $p = 2$:



$$= \tfrac{1}{2}(-i\lambda)^2\, \overline{\phi_1\phi_{y_1}}\; \overline{\phi_2\phi_{y_1}}\; \overline{\phi_3\phi_{y_2}}\; \left(\overline{\phi_{y_1}\phi_{y_2}}\right)^2$$

The combinatorial factor is obtained as follows. There are 4 ways of contracting ϕ_1 with $\phi_{y_1}^4$, which leaves 3 ways of contracting ϕ_2 with $\phi_{y_1}^3$. Similarly, here are 4 ways of contracting ϕ_3 with $\phi_{y_2}^4$, which leaves 3 ways of contracting ϕ_4 with $\phi_{y_2}^3$. Finally we have to contract $\phi_{y_1}^2$ with $\phi_{y_2}^2$, and there are 2 ways of doing this. We have

$$\left(\frac{1}{4!}\right)^2 \times 4 \cdot 3 \cdot 4 \cdot 3 \cdot 2 = \frac{1}{2}$$

As an exercise you should verify there are five other connected diagrams for the $p = 2$ case and they all have the same combinatorial factor of $\frac{1}{2}$.

5.4 Feynman Graphs

The diagrammatic language above is very useful in computing Green's functions in perturbation theory. But often we are interested in scattering amplitudes which are obtained from the Fourier transform $\widetilde{G}^{(n)}(k_1, \ldots, k_N)$ by the LSZ reduction formula. So it is convenient to replace the rules for the diagrammatic analysis above so that one obtains directly the Fourier transforms $\widetilde{G}^{(n)}$ or even the corresponding scattering amplitude $i\mathcal{A}$.

To this effect, in computing (5.9) we use

$$\overline{\phi_a\phi_b} = \int \frac{d^4q}{(2\pi)^4}\, e^{iq\cdot(x_a - x_b)} \frac{i}{q^2 - m^2 + i\epsilon}\,. \tag{5.10}$$

For perturbation theory the interaction term, $-i\lambda\int d^4y\,\phi^4(y)$, is as in (5.9) but integrated over space-time. Each interaction term gives a factor of

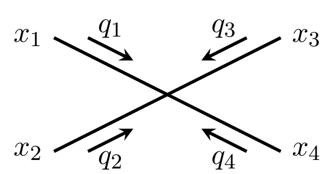
$$-i\lambda\int d^4y\,e^{-iy\cdot\sum_n q_n} = -i\lambda(2\pi)^4\delta^{(4)}(\sum_n q_n)\,,$$

where the q_n are from the four contractions $\prod_{i=1}^4 \overbrace{\phi(y)\phi}(x_i)$. To keep track of the signs $\pm q\cdot x$ in the arguments of the exponential, it is convenient to think of q as an arrow: in (5.10) it is directed from x_a to x_b.

So we have new rules:

1. Draw diagrams with n external "legs" (all topologically distinct diagrams).
2. For each topology assign momenta q_i to each line, including external legs. The assignment is directional: draw an arrow to indicate the direction of q_i, arbitrarily.
3. Every external line carries a factor
$$\int\frac{d^4q_n}{(2\pi)^4}e^{\pm iq_n\cdot x_n}\frac{i}{q_n^2-m^2+i\epsilon}\,.$$
with the plus sign if the arrow for q_n is drawn pointing into the diagram, minus if it points out.
4. Every internal line carries a factor
$$\int\frac{d^4q_i}{(2\pi)^4}\frac{i}{q_i^2-m^2+i\epsilon}\,.$$
5. Each vertex carries a factor
$$-i\lambda(2\pi)^4\delta^{(4)}(\sum_n(\pm)q_n)$$
where q_n are the momenta of the lines at the vertex, with the sign assignment $+1$ if q_n is directed into the vertex and -1 if directed out of the vertex.
6. Introduce a correction symmetry factor, as before.

For example, the connected diagram,



gives

$$\prod_{i=1}^{4}\int \frac{d^4q_n}{(2\pi)^4}e^{iq_n\cdot x_n}\frac{i}{q_n^2-m^2+i\epsilon}(-i\lambda)(2\pi)^4\delta^{(4)}(\sum_{n=1}^{4}q_n)$$

If the contraction involves an external leg, when taking the Fourier transform the corresponding coordinate x_i is integrated, $\int d^4x_i e^{-ik_i\cdot x_i}$. This gives

$$\int d^4x_i e^{-ik_i\cdot x_i}\int \frac{d^4q}{(2\pi)^4}, e^{iq\cdot(x_i-y)}\frac{i}{q^2-m^2+i\epsilon} = e^{-ik_i\cdot y}\frac{i}{k_i^2-m^2+i\epsilon}$$

Now recall, Eq. (5.5), that $\widetilde{G}^{(n)}(k_1,\ldots,k_n)$ is not really the Fourier transform of $G^{(n)}(x_1,\ldots,x_n)$, but rather

$$\int\prod_{i=1}^{n}d^4x_i\, e^{-i\sum_{i=1}^{n}k_i\cdot x_i}G^{(n)}(x_1,\ldots,x_n) = (2\pi)^4\delta^{(4)}(\sum_i k_i)\widetilde{G}^{(n)}(k_1,\ldots,k_n)\,.$$

The *Feynman rules* tell us how to compute for $(2\pi)^4\delta^{(4)}(\sum_i k_i)\widetilde{G}^{(n)}(k_1,\ldots,k_n)$ in perturbation theory:

1. Draw diagrams with n external "legs" (all topologically distinct diagrams).

2. For each topology find the inequivalent ways of assigning momenta k_i to each external leg. The assignment is directional: k_i goes into the diagram, "out" to "in" if $k_i^0 > 0$ (draw an arrow to indicate this).

3. Assign a momentum q_n, $n = 1,\ldots,I$ to each internal line. Draw an arrow to indicate this momentum direction, arbitrarily.

4. Every external line carries a factor

$$\frac{i}{k_i^2-m^2+i\epsilon}\,.$$

5. Every internal line carries a factor

$$\int\frac{d^4q_i}{(2\pi)^4}\frac{i}{q_i^2-m^2+i\epsilon}\,.$$

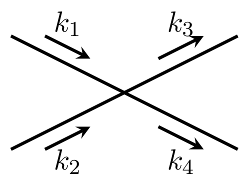
6. Each vertex carries a factor

$$-i\lambda(2\pi)^4\delta^{(4)}(\sum_n(\pm)p_n)$$

where p_n are the momenta of the lines at the vertex, with the sign assignment $+1$ if p_n is directed into the vertex and -1 if directed out of the vertex.

7. Introduce a correction symmetry factor, as before.

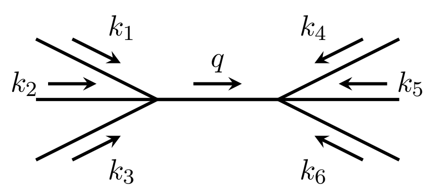
So, for example, to order λ in the perturbative expansion, the connected diagram for the 4-point function is



corresponding to

$$\widetilde{G}_1^{(4)}(k_1,\ldots,k_4) = \prod_{n=1}^{4} \frac{i}{k_n^2 - m^2 + i\epsilon} (-i\lambda)(2\pi)^4\delta^{(4)}(\sum_{n=1}^{4} k_n)$$

Here is another example, a contribution to order λ^2 to $\widetilde{G}^{(6)}$:



corresponding to

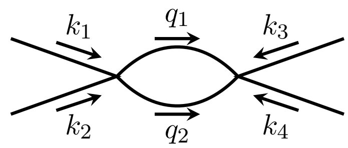
$$\prod_{n=1}^{6}\left(\frac{i}{k_n^2 - m^2 + i\epsilon}\right)\int \frac{d^4q}{(2\pi)^4}\frac{i}{q^2 - m^2 + i\epsilon}$$
$$\times \left[(-i\lambda)(2\pi)^4\delta^{(4)}(\sum_{n=1}^{3} k_n - q)\right]\left[(-i\lambda)(2\pi)^4\delta^{(4)}(\sum_{n=3}^{6} k_n + q)\right]$$
$$= (2\pi)^4\delta^{(4)}(\sum_{n=1}^{6} k_n)\left[-\lambda^2\prod_{n=1}^{6}\left(\frac{i}{k_n^2 - m^2}\right)\frac{i}{(k_1+k_2+k_3)^2 - m^2}\right]$$

leading to

$$\widetilde{G}_{\mathrm{conn}}^{(6)}(k_1,\ldots,k_6) = -\lambda^2\prod_{n=1}^{6}\left(\frac{i}{k_n^2 - m^2}\right)\frac{i}{(k_1+k_2+k_3)^2 - m^2} + \cdots$$

where the ellipses stand for other connected graphs at order λ^2 (can you display them?) plus terms of higher order in λ. We have removed the $i\epsilon$ from the propagators in the last line since all integrals have been performed.

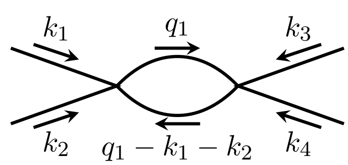
One more example:



gives

$$
\begin{aligned}
&\frac{1}{2}\prod_{n=1}^{4}\left(\frac{i}{k_n^2-m^2}\right)\prod_{i=1}^{2}\left(\int\frac{d^4q_i}{(2\pi)^4}\,\frac{i}{q_i^2-m^2+i\epsilon}\right)\\
&\quad\times\Big[(-i\lambda)(2\pi)^4\delta^{(4)}(k_1+k_2-q_1-q_2)\Big]\Big[(-i\lambda)(2\pi)^4\delta^{(4)}(k_3+k_4+q_1+q_1)\Big]\\
&=(2\pi)^4\delta^{(4)}(\sum_{n=1}^{4}k_n)\frac{1}{2}(-i\lambda)^2\prod_{n=1}^{4}\left(\frac{i}{k_n^2-m^2}\right)\int\frac{d^4q_1}{(2\pi)^4}\,\frac{i}{q_1^2-m^2+i\epsilon}\,\frac{i}{(k_1+k_2-q_1)^2-m^2+i\epsilon}
\end{aligned}
$$

Notice that this result involves a non-trivial integration. This occurs in any diagram for which there is a closed circuit of internal lines: momentum conservation at each vertex, enforced by δ-functions, does not completely fix the momentum of the internal lines. In this example the momentum q_1 appears in two propagators, tracing a closed trajectory, a *loop*. The situation is depicted in a new type of diagram in which the delta functions of momentum conservation have been explicitly accounted for (except an single factor that gives momentum conservation of the external momentum):

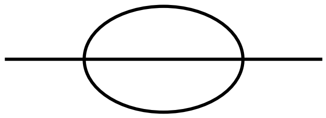


There is no longer an integration for each internal line. Instead there as in integral only over the undetermined momentum q_1. We call q_1 a *loop momentum* and $\int d^4q$ a *loop integral*.

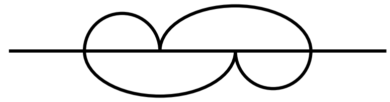
This generalizes. It is always the case that that the δ-functions at each vertex impose momentum conservation and therefore one can always recast the product of *delta*-functions as one $\delta^{(4)}(\sum_n k_n)$ of external momentum times the remaining delta functions. The number L of loop integrals we are left to do in any given diagram is the number of internal lines I minus the number of *delta*-functions, taking away the one for overall momentum conservation. If there are V vertices, we then have $V-1$ δ-functions and therefore

$$L = I - V + 1$$

In our example above, $I = 2$, $V = 2$ and we had $L = 2 - 2 + 1 = 1$ loops. Here are few more examples:

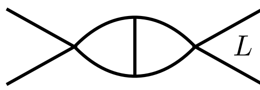


$L = 3 - 2 + 1 = 2$



$L = 7 - 4 + 1 = 4$

We can also have a theory with 3-point and 4-point vertices, as in $\frac{1}{3!} g\phi^3 + \frac{1}{4!}\lambda\phi^4$; here is an example:



$L = 5 - 4 + 2 = 2$

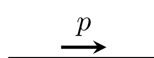
This suggest a more compact set of *Feynman rules* to compute $\widetilde{G}^{(n)}(k_1, \ldots, k_n)$:

1. Draw diagrams with n external "legs" (all topologically distinct diagrams).

2. For each topology find the inequivalent ways of assigning momenta k_i to each external leg. The assignment is directional: k_i goes into the diagram. Draw an arrow to indicate this.

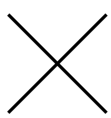
3. Assign q_i, $i = 1, \ldots, L$ momenta to internal lines; draw arrow indicating direction. Assign momenta to remaining $I - L = V - 1$ internal lines by enforcing momentum conservation: at each vertex $\sum p_{\text{in}} = \sum p_{\text{out}}$.

4. For each line



$$= \frac{i}{p^2 - m^2 + i\epsilon}$$

5. For each vertex,



$$= -i\lambda$$

6. Integrate:

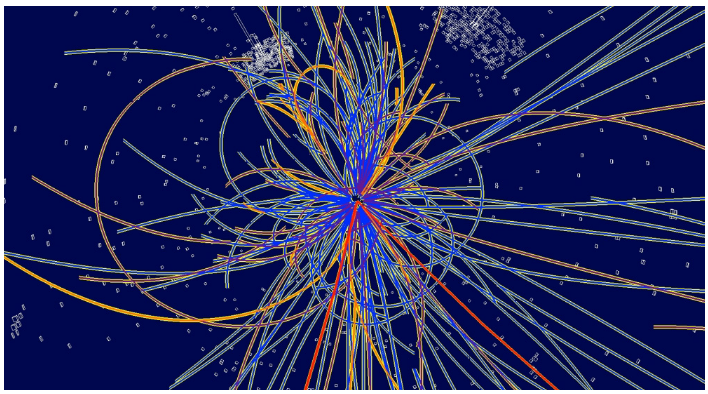
$$\prod_{n=1}^{L} \int \frac{d^4 q_n}{(2\pi)^4}$$

.

7. Symmetry factor $1/S$ as needed.

5.5 Cross Section

In a common experimental setup two beams of elementary particles are accelerated in opposite directions and brought into face to face encounter. Some fraction of the particles in the beams collide. The collision results in a spray of elementary particles emanating form the collision point and an array of detectors surrounding the area register these outgoing particles. Here as a computer generated image of the tracks made by charged particles that are sprayed out of the head-to-head collision of two protons, projected onto a plane transverse to the direction of the protons:



In another common setup a beam of particle impinges on a collection of stationary targets. This second set-up is, of course, just the first one as seen by an observer at rest with the second "beam." We say this observer is in the "lab frame."

We are after a measure of how likely are these collisions to occur. The cross section, σ, for scattering is defined through

$$\frac{\text{number of collisions}}{\text{unit time}} = (\text{flux}) \times \sigma\,.$$

To calculate, rather than computing the number of collisions per unit time from the actual flux, we use unit flux (that of one-on-one particles) and therefore

$$\frac{\text{collision probability}}{\text{unit time}} = (\text{unit flux}) \times \sigma\,.$$

We have an initial state $|i\rangle$ that consists of two particles, a final state $|f\rangle$ that consists of n-particles ($n/ge2$). The probability that $|i\rangle$ evolves into $|f\rangle$ is

$$P_{i\to f} = |\langle f\,\text{out}|i\,\text{in}\rangle|^2 = |\langle f\,\text{in}|S|i\,\text{in}\rangle|^2 = |\langle f|S-1|i\rangle|^2$$

where in the last step we ignore the $f = i$ case (no collision, hence subtract 1 from S) and suppressed the "in" label (we will get tired of carrying it around).

We have to be careful to ask the right question: since we have continuum normalization of states, if we are overly selective in what we want for $|f\rangle$, the probability of finding it in $|i\rangle$ will vanish. Recall if you drop a pin on a piece of paper the probability of hitting a given point on the paper, say, (x_0, y_0), is zero, since a point is a set of measure zero in the set of points that comprise the area of the paper. Likewise, if we set $|f\rangle = |\vec{k}_1, \ldots, \vec{k}_n\rangle$ we'll find $P_{i\to f} = 0$. Instead we project out a subspace of $\mathcal{F}$, rather than a single state. Instead of

$$|\langle f|S-1|i\rangle|^2 = \langle i|(S-1)^\dagger|f\rangle\langle f|S-1|i\rangle$$

we take

$$\langle i|(S-1)^\dagger\Big(\sum_{\substack{f\\ \text{some states}}}|f\rangle\langle f|\Big)(S-1)|i\rangle$$

In particular, for n particles in the final state we have

$$\sum_f |f\rangle\langle f| \to \int (dk_1)\cdots(dk_n)|\vec{k}_1, \ldots, \vec{k}_n\rangle\langle\vec{k}_1, \ldots, \vec{k}_n|$$

where

- we may not want to sum over all possible momenta, so the integrals can be restricted
- must avoid double counting from indistinguishable particles

Suppose particles 1 and 2 are indistinguishable (but the rest are not). Then to avoid double counting one should write

$$\sum_f |f\rangle\langle f| \to \frac{1}{2}\int (dk_1)\cdots(dk_n)|\vec{k}_1, \ldots, \vec{k}_n\rangle\langle\vec{k}_1, \ldots, \vec{k}_n|$$

If 1,2,3 are indistinguishable then the pre-factor becomes 1/3! since the order of $\vec{k}_1$, $\vec{k}_2$ and $\vec{k}_3$ in the label of the state is immaterial. More generally,

$$\sum_f |f\rangle\langle f| \to \frac{1}{S}\prod_{i=1}^{n}(dk_i)\,|\vec{k}_1, \ldots, \vec{k}_n\rangle\langle\vec{k}_1, \ldots, \vec{k}_n|$$

where $S = m_1!m_2!\cdots$ where m_i is the number of identical particles of type i ($\sum_i m_i = n$).

We are ready to give a probability:

$$P_{i\to f} = \frac{\langle i|(S-1)^\dagger \sum |f\rangle\langle f|S-1|i\rangle}{\langle i|i\rangle}$$

We have divided by $\langle i|i\rangle$ because states must be normalized for proper interpretation. Did not divide by normalization of $|f\rangle$ because it is included properly in relativistic measure in the sum over states. We next recast this in terms of the scattering amplitude,

$$\langle f|S-1|i\rangle = (2\pi)^4\delta^{(4)}(P_f - P_i)\, i\mathcal{A}(i \to f).$$

At this point we choose an initial state of plane waves with definite momentum, $|i\rangle = |\vec{p}_1, \vec{p}_2\rangle$. We have

$$\langle i|i\rangle = \langle \vec{p}_1, \vec{p}_2|\vec{p}_1, \vec{p}_2\rangle = \langle \vec{p}_1|\vec{p}_1\rangle\langle \vec{p}_2|\vec{p}_2\rangle$$

but since in general $\langle \vec{p}|\vec{k}\rangle = 2E_{\vec{k}}(2\pi)^3\delta^{(3)}(\vec{p}-\vec{k})$, we have, $\langle \vec{p}_1|\vec{p}_1\rangle = 2E_{\vec{k}}(2\pi)^3\delta^{(3)}(0)$. This is embarrassing, but not disastrous. There are two ways of dealing with this problem. It is not very hard to use wave-packets, which can be properly normalized, instead of plane waves, but won't do here; check it out in some of the textbooks in our bibliography. We have an alternative means of dealing with this problem, which is by putting the system in a finite box of volume V. We have already done so in computing phase space. We discovered there that the correct interpretation of this infinity is $(2\pi)^3\delta^{(3)}(0) \to V$. In fact, we will also use, more generally, $(2\pi)^4\delta^{(4)}(0) \to VT$, where $T = t_{\text{f}inal} - t_{initial}$. So we can write $\langle i|i\rangle = 4E_1E_2V^2$. Similarly

$$\begin{aligned}
|\langle f|S-1|i\rangle|^2 &= \Big((2\pi)^4\delta^{(4)}(P_f - P_i)\Big)^2 |\mathcal{A}(i\to f)|^2 \\
&= \Big((2\pi)^4\delta^{(4)}(0)\Big)(2\pi)^4\delta^{(4)}(P_f - P_i)|\mathcal{A}(i\to f)|^2 \\
&= VT(2\pi)^4\delta^{(4)}(P_f - P_i)|\mathcal{A}(i\to f)|^2
\end{aligned}$$

Putting it all together:

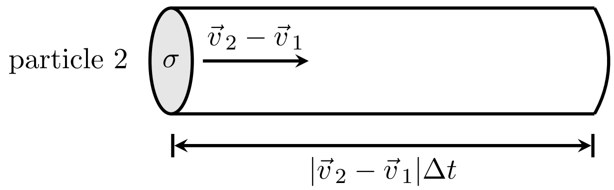
$$\begin{aligned}
\frac{\text{probability}}{\text{time}} = \frac{1}{T}\frac{VT\frac{1}{S}\int\prod_i(dk_i)\,(2\pi)^4\delta^{(4)}(P_f - P_i)|\mathcal{A}(i\to f)|^2}{4E_1E_2V^2} \\
= \frac{1}{V}\frac{1}{4E_1E_2}\int|\mathcal{A}(i\to f)|^2 \mathrm{d}\Phi_n
\end{aligned}$$

where

$$\mathrm{d}\Phi_n = \frac{1}{S}\,(2\pi)^4\delta^{(4)}(P_f - P_i)\,(dk_1)\cdots(dk_n)$$

is the Lorentz-invariant n-particle phase space.

Finally, in order to determine the cross section σ we need to divide the above by the unit flux. Assume particle 1 is uniformly distributed in a box of volume V. The probability of finding it is a sub-volume v is v/V. We want v to be the interaction volume, so project a volume forward of particle 2, in the direction of the relative motion $\vec{v}_2 - \vec{v}_1$, with cross sectional area σ perpendicular to that direction, when particle 2 moves over a time Δt, as in the following figure:



This has

$$\frac{\text{volume}}{V} = \frac{(|\vec{v}_2 - \vec{v}_1|\Delta t)\sigma}{V} \quad \Rightarrow \quad \frac{\text{probability}}{\text{unit time}} = \frac{|\vec{v}_2 - \vec{v}_1|\sigma}{V}$$

Comparing with the above probability per unit time we have

$$\boxed{\mathrm{d}\sigma = \frac{1}{4E_1E_2|\vec{v}_2 - \vec{v}_1|}\,|\mathcal{A}|^2\mathrm{d}\Phi}$$

where we have written $\mathrm{d}\sigma$ rather than σ to remind us that $\mathrm{d}\Phi$ will be integrated over: $\sigma = \int \mathrm{d}\sigma = \frac{1}{4E_1E_2|\vec{v}_2-\vec{v}_1|}\int |\mathcal{A}|^2\mathrm{d}\Phi$.

The factor $E_1E_2|\vec{v}_2-\vec{v}_1|$ is invariant under boosts along the direction of $\vec{v}_2-\vec{v}_1$. This is most easily seen in a frame where $\vec{v}_1$ and $\vec{v}_2$ are along the z-axis. Then

$$E_1E_2|\vec{v}_2 - \vec{v}_1| = E_1E_2|\frac{p_2}{E_2} - \frac{p_1}{E_1} = |p_2E_1 - p_1E_2| = |\epsilon_{12\mu\nu}p_1^\mu p_2^\nu|$$

is invariant to boosts in the 3-direction. It is useful to compute this factor in the two most common frames, once and for all. In the *Lab frame*: $\vec{p}_2 = 0$, $E_2 = m_2$ so $E_1E_2|\vec{v}_2 - \vec{v}_1| = |p_2E_1 - p_1E_2| = m_2p_1$. In the center of mass, or *CM frame*, $\vec{p}_2 + \vec{p}_1 = 0$, so that $\vec{p}_1 = -\vec{p}_2 \equiv \vec{p}$ and $E_1E_2|\vec{v}_2 - \vec{v}_1| = |p_2E_1 - p_1E_2| = |\vec{p}|\sqrt{(p_1+p_2)^2}$ where $p_{1,2}$ are 4-vectors. Let

$$s \equiv (p_1 + p_2)^2$$

so that $s = m_1^2 + m_2^2 + 2E_1E_2 + 2|\vec{p}|^2$. Since $E_i = \sqrt{|\vec{p}|^2 + m_i^2}$ we have an equation relating $|\vec{p}|^2$ to s, which we solve:

$$|\vec{p}|^2 = \frac{(s - m_1^2 - m_2^2)^2 - 4m_2^2m_2^2}{4s}$$

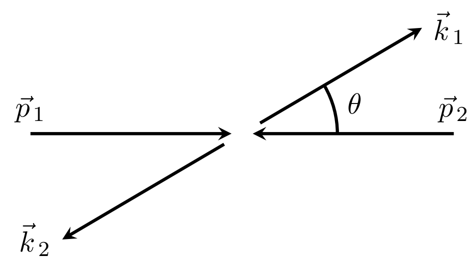
This gives

$$4E_1E_2|\vec{v}_2 - \vec{v}_1| = |\vec{p}|\sqrt{s} = 2\sqrt{(s - m_1^2 - m_2^2)^2 - 4m_1^2m_2^2}\,.$$

The last expression is valid in any frame boosted along $\vec{v}_2 - \vec{v}_1$, and we can write

$$\mathrm{d}\sigma = \frac{1}{2\sqrt{(s - m_1^2 - m_2^2)^2 - 4m_1^2m_2^2}}\,|\mathcal{A}|^2\mathrm{d}\Phi$$

Example: $2 \to 2$ **scattering, identical particles**



We have for identical particles $m_1 = m_2 = m$

$$\begin{aligned}
\mathrm{d}\Phi_2 &= \frac{d^3k_1}{(2\pi)^3 2E_1}\frac{d^3k_2}{(2\pi)^3 2E_2}\,(2\pi)^4\delta^{(4)}(P-k_1k_2) \\
&= \frac{d^4k_1\,d^4k_2}{(2\pi)^2}\theta(k_1^0)\theta(k_2^0)\delta(k_1^2-m^2)\delta(k_2^2-m^2)\delta^{(4)}(P-k_1k_2)
\end{aligned}$$

Change variables:

$$\begin{aligned} p &= k_1 + k_2 \\ q &= \tfrac{1}{2}(k_1 - k_2) \end{aligned} \qquad \Leftrightarrow \qquad \begin{aligned} k_1 &= \tfrac{1}{2}p + q \\ k_2 &= \tfrac{1}{2}p - q \end{aligned}$$

Note that the Jacobian of the transformation, $\left|\frac{\partial(k_1,k_2)}{\partial(p,q)}\right| = 1$. Then,

$$\begin{aligned}
\mathrm{d}\Phi_2 &= \frac{1}{(2\pi)^2}d^4p\delta^{(4)}(P-p)d^4q\theta(\tfrac{1}{2}p^0+q^0)\theta(\tfrac{1}{2}p^0-q^0)\delta((\tfrac{1}{2}p+q)^2-m^2)\delta((\tfrac{1}{2}p-q)^2-m^2) \\
&= \frac{1}{(2\pi)^2}d^4q\,\theta(\tfrac{1}{2}P^0-|q^0|)\delta(\tfrac{1}{4}P^2+P\cdot q+q^2-m^2)\delta(2P\cdot q)
\end{aligned}$$

In the CM frame, $\vec{P} = 0$, this is simple:

$$\begin{aligned}
\mathrm{d}\Phi_2 &= \frac{1}{(2\pi)^2}dq^0|\vec{q}|^2 d|\vec{q}|d\cos\theta d\phi\,\theta(\tfrac{1}{2}P^0-|q^0|)\delta(\tfrac{1}{4}P^2+(q^0)^2-|\vec{q}|^2-m^2)\tfrac{1}{2P^0}\delta(q^0) \\
&= \frac{1}{(2\pi)^2}d\cos\theta d\phi\,\frac{1}{2P^0}\frac{1}{2}\sqrt{\tfrac{1}{4}(P^0)^2-m^2} \\
&= \frac{1}{8(2\pi)^2}d\cos\theta d\phi\,\sqrt{1-4m^2/s}
\end{aligned}$$

where $s = P^2 = (p_1+p_2)^2$ as before. Therefore

$$\begin{aligned}
\frac{\mathrm{d}\sigma}{d\cos\theta d\phi} &= \frac{1}{2}\frac{1}{8(2\pi)^2}\sqrt{1-4m^2/s}\frac{1}{2\sqrt{(s-2m^2)^2-4m^2}}|\mathcal{A}|^2 \\
&= \frac{1}{32}\frac{1}{(2\pi)^2}\frac{1}{s}|\mathcal{A}|^2
\end{aligned}$$

Often $\mathcal{A}$ is independent of ϕ, so

$$\frac{d\sigma}{d\cos\theta} = \frac{1}{64\pi}\frac{1}{s}|\mathcal{A}|^2$$

5.6 LSZ reduction, again

We want to establish the LSZ reduction formula. We don;t pretend to give a complete proof. The objective is to express the S-matrix in terms of Green' s functions, vacuum expectation values of time ordered products, of in fields. Consider

$${}_{\text{out}}\langle\psi|\vec{k}_1, \vec{k}_2, \ldots, \vec{k}_n\rangle_{\text{in}} = {}_{\text{out}}\langle\psi|\vec{k}_1\chi\rangle_{\text{in}} = {}_{\text{out}}\langle\psi|\alpha^\dagger_{\vec{k}_1\text{in}}\chi\rangle_{\text{in}}$$

Recall

$$\begin{aligned}\phi_{\text{in}}(x) &= \int (dq)\left(\alpha_{\vec{q}\text{in}}e^{-iq\cdot x} + \alpha^\dagger_{\vec{q}\text{in}}e^{iq\cdot x}\right)\\ \partial_t\phi_{\text{in}}(x) &= \int (dq)\left(-iE_{\vec{q}}\alpha_{\vec{q}\text{in}}e^{-iq\cdot x} + iE_{\vec{q}}\alpha^\dagger_{\vec{q}\text{in}}e^{iq\cdot x}\right)\end{aligned}$$

Inverting these,

$$\alpha^\dagger_{\vec{q}\text{in}} = -i\int d^3x\, e^{-iq\cdot x}\overleftrightarrow{\partial_t}\,\phi_{\text{in}}(x)$$

So we have

$$\begin{aligned}{}_{\text{out}}\langle\psi|\vec{k}_1\chi\rangle_{\text{in}} &= -i\int d^3x\, e^{-ik_1\cdot x}\overleftrightarrow{\partial_t}\;{}_{\text{out}}\langle\psi|\phi_{\text{in}}(x)|\chi\rangle_{\text{in}}\\ &= -iZ^{-\frac{1}{2}}\int d^3x\, e^{-ik_1\cdot x}\overleftrightarrow{\partial_t}\;{}_{\text{out}}\langle\psi|\phi(x)|\chi\rangle_{\text{in}} \quad \text{as } t\to-\infty.\end{aligned}$$

Now we use the fundamental theorem of calculus, $g(t_2) = g(t_1) + \int_{t_1}^{t_2} dt\frac{dg}{dt}$, with $t_2 \to \infty$ and $t_1 \to -\infty$ so that

$$\begin{aligned}\lim_{t\to-\infty}\int d^3x\, e^{-ik_1\cdot x}\overleftrightarrow{\partial_t}\;{}_{\text{out}}\langle\psi|\phi(x)|\chi\rangle_{\text{in}} &= \lim_{t\to\infty}\int d^3x\, e^{-ik_1\cdot x}\overleftrightarrow{\partial_t}\;{}_{\text{out}}\langle\psi|\phi(x)|\chi\rangle_{\text{in}}\\ &\quad -\int d^4x\,\partial_t\left(e^{-ik_1\cdot x}\overleftrightarrow{\partial_t}\;{}_{\text{out}}\langle\psi|\phi(x)|\chi\rangle_{\text{in}}\right)\end{aligned}$$

The first term on the right hand side times $-iZ^{-\frac{1}{2}}$ is ${}_{\text{out}}\langle\psi|\alpha^\dagger_{\vec{k}_1\text{out}}|\chi\rangle_{\text{in}}$, as can be seen by reversing the steps. If $|\psi\rangle_{\text{out}} = |\vec{p}_1, \vec{p}_2, \ldots, \vec{p}_{n'}\rangle_{\text{out}}$ then $\alpha^\dagger_{\vec{k}_1\text{out}}|\psi\rangle_{\text{out}} = (2\pi)^3 2E_{\vec{k}_1}\delta^{(3)}(\vec{k}_1-\vec{p}_1)|\vec{p}_2, \ldots, \vec{p}_{n'}\rangle_{\text{out}}+\cdots+(2\pi)^3 2E_{\vec{k}_1}\delta^{(3)}(\vec{k}_1-\vec{p}_{n'})|\vec{p}_1, \ldots, \vec{p}_{n'-1}\rangle_{\text{out}}$ corresponds to a particle not participating in the scattering. We have no use for this. For the second term on the right hand side we use $\partial_t(f(t)\overleftrightarrow{\partial_t}\, g(t)) = f\partial_t^2 g - (\partial_t^2 f)g$ so it is

$$-\int d^4x\, e^{-iq\cdot x}(\partial_t^2 + E_{\vec q}^2)\,{}_{\rm out}\langle\psi|\phi(x)|\chi\rangle_{\rm in}$$

Using $E_{\vec q}^2 e^{-iq\cdot x} = (|\vec q|^2 + m^2)e^{-iq\cdot x} = (-\vec\nabla^2 + m^2)e^{-iq\cdot x}$ and integrating by parts we have finally

$${}_{\rm out}\langle\psi|\vec k_1\chi\rangle_{\rm in} = {}_{\rm out}\langle\psi|\alpha^\dagger_{\vec k_1 \rm out}|\chi\rangle_{\rm in} + iZ^{-\frac12}\int d^4x\, e^{-ik_1\cdot x}(\partial^2 + m^2)\,{}_{\rm out}\langle\psi|\phi(x)|\chi\rangle_{\rm in}$$

Similarly

$${}_{\rm out}\langle\psi\vec p|\chi\rangle_{\rm in} = {}_{\rm out}\langle\psi|\alpha^\dagger_{\vec p \rm in}|\chi\rangle_{\rm in} + iZ^{-\frac12}\int d^4x\, e^{ip\cdot x}(\partial^2 + m^2)\,{}_{\rm out}\langle\psi|\phi(x)|\chi\rangle_{\rm in} \quad (5.11)$$

We woud like to repeat the process until we remove all particles from $|\psi\rangle_{\rm out}$ and $|\chi\rangle_{\rm in}$. To see how this goes move a particle off from $|\chi\rangle_{\rm in}$ from what we already had:

$$\begin{aligned}
{}_{\rm out}\langle\psi|\phi(0)|\vec k\,\chi'\rangle_{\rm in} = {}_{\rm out}\langle\psi|\phi(0)\alpha^\dagger_{\vec k \rm in}|\chi'\rangle_{\rm in} &= -iZ^{-\frac12}\lim_{x^0\to-\infty}\int d^3x\, e^{-ik\cdot x}\overleftrightarrow{\partial}_{x^0}\,{}_{\rm out}\langle\psi|\phi(0)\phi(x)|\chi'\rangle_{\rm in}\\
&= iZ^{-\frac12}\int d^4x\,\partial_{x^0}\left(e^{-ik\cdot x}\overleftrightarrow{\partial}_{x^0}\,{}_{\rm out}\langle\psi|\phi(0)\phi(x)|\chi'\rangle_{\rm in}\right)\\
&\quad - iZ^{-\frac12}\lim_{x^0\to\infty}\int d^3x\, e^{-ik\cdot x}\overleftrightarrow{\partial}_{x^0}\,{}_{\rm out}\langle\psi|\phi(0)\phi(x)|\chi'\rangle_{\rm in} \qquad (5.12)
\end{aligned}$$

In the last expression we would like to move $\phi(x)$ to the left of $\phi(0)$ so that we may turn $\phi(x)$ as $x^0\to\infty$ into $\alpha^\dagger_{\vec k \rm out}$ acting on $\langle{\rm out}|$. To this end we rewrite the first term in the last expression in (5.12) using

$$\begin{aligned}
\phi(0)\phi(x) &= (\theta(-x^0) + \theta(x^0))\phi(0)\phi(x) + \theta(x^0)(\phi(x)\phi(0) - \phi(x)\phi(0))\\
&= T(\phi(x)\phi(0)) + \theta(x^0)[\phi(0),\phi(x)]
\end{aligned}$$

Then in

$$\int d^4x\,\partial_{x^0}\left(e^{-ik\cdot x}\overleftrightarrow{\partial}_{x^0}\,{}_{\rm out}\langle\psi|\theta(x^0)[\phi(0),\phi(x)]|\chi'\rangle_{\rm in}\right) \qquad (5.13)$$

when $\overleftrightarrow{\partial}_{x^0}$ hits $\theta(x^0)$ we get the equal time commutator $[\phi(0),\phi(x)] = 0$. So we have (5.13) is

$$\begin{aligned}
&= \int d^4x\,\partial_{x^0}\left[\theta(x^0)\left(e^{-ik\cdot x}\overleftrightarrow{\partial}_{x^0}\,{}_{\rm out}\langle\psi|[\phi(0),\phi(x)]|\chi'\rangle_{\rm in}\right)\right]\\
&\qquad = \lim_{x^0\to\infty}\int d^3x\, e^{-ik\cdot x}\overleftrightarrow{\partial}_{x^0}\,{}_{\rm out}\langle\psi|[\phi(0),\phi(x)]|\chi'\rangle_{\rm in}
\end{aligned}$$

Combining this with the last term in (5.12) gives precisely what we want: it reverses the order of $\phi(0)\phi(x)$ so that

$$ {}_{\rm out}\langle\psi|\phi(0)|\vec{k}\,\chi'\rangle_{\rm in} = iZ^{-\frac{1}{2}}\int d^4x\,\left(e^{-ik\cdot x}(\partial^2+m^2)\,{}_{\rm out}\langle\psi|T(\phi(0)\phi(x))|\chi'\rangle_{\rm in}\right)+{}_{\rm out}\langle\psi|\alpha^{\dagger}_{\vec{k}\,{\rm out}}\phi(0)|\chi'\rangle_{\rm in} $$

We thus arrive at

$$ \begin{aligned} {}_{\rm out}\langle\psi|\vec{k}_1\vec{k}_2\chi'\rangle_{\rm in} &= (iZ^{-\frac{1}{2}})^2\int d^4x_1 d^4x_2 e^{-ik_1\cdot x_1-ik_2\cdot x_2} \\ &\quad\times(\partial^2_{x_1}+m^2)(\partial^2_{x_2}+m^2)\,{}_{\rm out}\langle\psi|T(\phi(0)\phi(x))|\chi'\rangle_{\rm in}+\text{disconnected} \end{aligned} $$

One can repeat the process until all particles in $|{\rm in}\rangle$ are removed and we are left with $|0\rangle_{\rm in}=|0\rangle$. The argument above can be streamlined by replacing $T(\phi(0)\phi(x))$ for $\phi(0)\phi(x)$ in the line above (5.12)), and this indeed becomes very convenient in completing the argument for arbitrary number of particles.

Similarly we can remove 1-particle states from ${}_{\rm out}\langle\vec{k}\,\psi|$ using (5.11) repeatedly.

5.7 Perturbation theory, again

We now give a proof of (5.6) that gives us the basis for perturbation theory. Consider

$$ G^{(n)}(x_1,\ldots,x_n)=\langle 0|T(\phi(x_1)\cdots\phi(x_n))|0\rangle\,. $$

Take for definiteness $x^0_1\geq x^0_2\geq\cdots\geq x^0_n$. Recall $\phi(x)=U(t)^{-1}\phi_{\rm in}U(t)$, ans use this in each ϕ in the Green's function:

$$ G^{(n)}(x_1,\ldots,x_n)=\langle 0|U^{-1}(t_1)\phi(x_1)U(t_1)U^{-1}(t_2)\phi(x_2)U(t_2)\cdots U^{-1}(t_n)\phi(x_n)U(t_n)|0\rangle\,. $$

Let $U(t,t')=U(t)U^{-1}(t')$. This satisfies

$$ U(\infty,-\infty)=S \tag{5.14} $$

$$ U(t,-\infty)=U(t) \tag{5.15} $$

$$ U(t,t')U(t',t'')=U(t,t'') \tag{5.16} $$

Moreover,

$$ i\frac{\partial U(t,t')}{\partial t}=H'_{\rm in}U(t,t'),\qquad\text{with } U(t',t')=1\,. $$

This is the same equation satisfied by $U(t)$, but with a different boundary condition. So the solution is the same only with different limits of integration,

$$ U(t_f,t_i)=T\exp\left(-i\int_{t_i}^{t_f}dtH'_{\rm in}(t)\right)\,. $$

We now have

$$\begin{aligned}
G^{(n)} &= \langle 0|U^{-1}(\infty)U(\infty,t_1)\phi_{\rm in}(x_1)U(t_1,t_2)\phi_{\rm in}(x_2)U(t_2,t_3)\cdots U(t_{n-1},t_n)\phi_{\rm in}(x_n)U(t_n,-\infty)U(-\infty)|0\rangle \\
&= \langle 0|U^{-1}(\infty)T\left(U(\infty,t_1)\phi_{\rm in}(x_1)U(t_1,t_2)\phi_{\rm in}(x_2)U(t_2,t_3)\cdots U(t_{n-1},t_n)\phi_{\rm in}(x_n)U(t_n,-\infty)\right)|0\rangle \\
&= \langle 0|U^{-1}(\infty)T\left(U(\infty,t_1)U(t_1,t_2)U(t_2,t_3)\cdots U(t_{n-1},t_n)U(t_n,-\infty)\phi_{\rm in}(x_1)\phi_{\rm in}(x_2)\cdots\phi_{\rm in}(x_n)\right)|0\rangle \\
&= \langle 0|U^{-1}(\infty)T\left(U(\infty,-\infty)\phi_{\rm in}(x_1)\phi_{\rm in}(x_2)\cdots\phi_{\rm in}(x_n)\right)|0\rangle
\end{aligned}$$

Uniqueness of the vacuum means that $|0\rangle$, $|0\rangle_{\rm in}$, $|0\rangle_{\rm out}$, are equal up to a phase. Moreover, $U(\infty)|0\rangle = S|0\rangle$ must be $|0\rangle$ up to a phase. To see this note that S commutes with Poincare transformations, $U(a^\mu,\Lambda)SU^\dagger(a^\mu,\Lambda) = S$ and $|0\rangle$ is the unique state (up to a phase) that is left invariant by a Poincare transformation. Then $U(a^\mu,\Lambda)(S|0\rangle) = U(a^\mu,\Lambda)SU^\dagger(a^\mu,\Lambda)U(a^\mu,\Lambda)|0\rangle = S|0\rangle$ so $S|0\rangle$ is invariant and hence equal to $|0\rangle$ up to a phase. So we have $\langle 0|U^{-1}(\infty) = \langle 0|U^{-1}(\infty)|0\rangle\,\langle 0|$. Using this and

$$U(\infty) = U(\infty,-\infty) = T\exp\left(-i\int_{-\infty}^{\infty} dt\, H'_{\rm in}(t)\right)$$

we finally have

$$G^{(n)}(x_1,\ldots,x_n) = \frac{{}_{\rm in}\langle 0|T\left(\phi_{\rm in}(x_1)\cdots\phi_{\rm in}(x_n)e^{i\int d^4x\,\mathcal{L}'_{\rm in}}\right)|0\rangle_{\rm in}}{{}_{\rm in}\langle 0|T\left(e^{i\int d^4x\,\mathcal{L}'_{\rm in}}\right)|0\rangle_{\rm in}}$$

Note that we replaced $|0\rangle_{\rm in}$ for $|0\rangle$ since the phases in numerator and denominator cancel.

This is not what we set out to prove. It is better. The denominator corresponds to a set of graphs without external legs. These vacuum graphs can also appear in the numerator, just multiplying any graph with external legs. It is a simple exercise to check that the vacuum graphs in the numerator are cancelled by the graphs in the denominator.

5.7.1 Generating Function for Green's Functions

Let

$$Z[J] = \langle 0|Te^{i\int d^4x\, J(x)\phi(x)}|0\rangle\,.$$

Then

$$G^{(n)}(x_1,\ldots,x_n) = \frac{1}{i}\frac{\delta}{\delta J(x_1)}\cdots\frac{1}{i}\frac{\delta}{\delta J(x_n)}Z[J]\bigg|_{J=0}$$

and

$$Z[J] = \frac{{}_{\rm in}\langle 0|Te^{i\int d^4x(\mathcal{L}'_{\rm in}+J(x)\phi_{\rm in}(x))}|0\rangle_{\rm in}}{{}_{\rm in}\langle 0|T\left(e^{i\int d^4x\mathcal{L}'_{\rm in}}\right)|0\rangle_{\rm in}}$$

This is a convenient way of summarizing the results above for $G^{(n)}$, all n. Note also that

$$Z[J] = \sum_n \frac{i^n}{n!} \int d^4x_1 \cdots d^4x_n \, J(x_1) \cdots J(x_n) G^{(n)}(x_1, \ldots, x_n) \, .$$

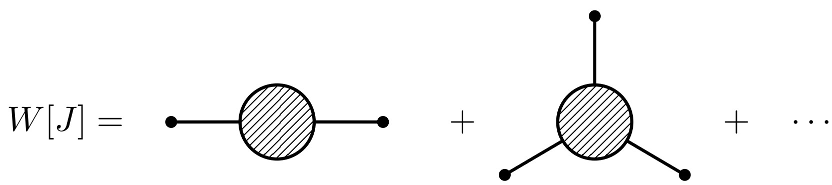
5.7.2 Generating Function for Connected Green's Functions

Similarly we define

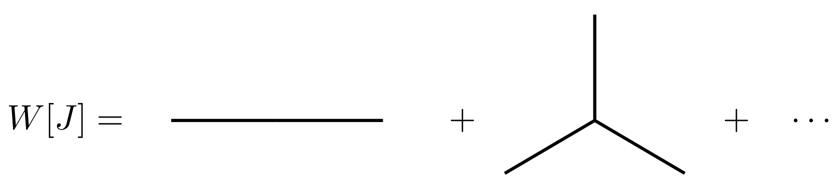
$$W[J] = \sum_n \frac{i^n}{n!} \int d^4x_1 \cdots d^4x_n \, J(x_1) \cdots J(x_n) G^{(n)}_{\rm conn}(x_1, \ldots, x_n) \, .$$

We will now show that $Z[J] = e^{iW[J]}$.

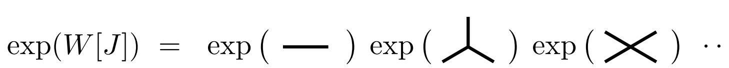
We use a diagrammatic notation to see how this works:



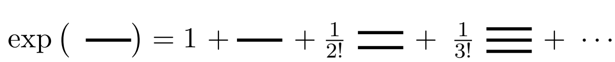
where the heavy dots ﬄ stand for $J(x)$, the hatch circles with n lines stand for $G^{(n)}$ and an integral $\frac{1}{n!}\prod_i \int d^4x_i$ in each term is understood. Let streamline notation for the purposes of this proof: remove the heavy dots (the ends of lines are understood as having them) and reduce the hatch circle to a point, so that the above figure is the same as



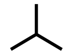
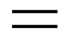
With this notation we consider the exponential of $W[J]$:



Now expand its exponential, as in



and reorganize by powers of J, that is, number fo external legs:

$\exp W[J] = 1 +$ $+$  $+\Big(\frac{1}{2!}$  $+$  $\Big) + \cdots$

where the ellipses stand for terms with five or more legs. Let's analyze in more detail the term in parenthesis: we want to show that it gives $G^{(4)}$ (times sources, 1/4! and integrated). The "cross" stands for

$$\frac{1}{4!}\int d^4y_1\cdots d^4y_4\, J(y_1)\cdots J(y_4)G^{(4)}_{\rm conn}(y_1,\ldots,y_4)$$

Now, $\frac{\delta^4}{\delta J(x_1)\cdots\delta J(x_4)}$ of this gives $G^{(4)}_{\rm conn}(x_1,\ldots,x_4)$. Note that the 4! is absent since there are 4! terms from the integral (same as in $\frac{d^4}{dx^4}x^4 = 4!$). Turning to the other term, the disconnected graph, we have

$$\frac{\delta^4}{\delta J(x_1)\cdots\delta J(x_4)}\frac{1}{2!}\left(\frac{1}{2!}\int d^4y d^4z\, J(y)J(z)G^{(2)}(y,z)\right)^2$$
$$= G^{(2)}(x_1,x_2)G^{(2)}(x_3,x_4) + G^{(2)}(x_1,x_3)G^{(2)}(x_2,x_4) + G^{(2)}(x_1,x_4)G^{(2)}(x_2,x_3)$$

If we take the for J-functional derivatives and set $J = 0$ these are the only terms we pick up in the expansion, so we have

$$\frac{\delta^4}{\delta J(x_1)\cdots\delta J(x_4)}e^{W[J]}\Big|_{J=0} = G^{(4)}_{\rm conn}(x_1,\ldots,x_4)$$
$$+\, G^{(2)}(x_1,x_2)G^{(2)}(x_3,x_4) + G^{(2)}(x_1,x_3)G^{(2)}(x_2,x_4) + G^{(2)}(x_1,x_4)G^{(2)}(x_2,x_3)$$
$$= G^{(4)}(x_1,\ldots,x_4) = \frac{\delta^4}{\delta J(x_1)\cdots\delta J(x_4)}Z[J]\Big|_{J=0}$$

In the general case, the term with J^n in $e^{W[J]}$ is a sum of all possible contributions of the form

$$\frac{1}{n_2!}\left(\frac{1}{2!}\int J_1J_2G^{(2)}_c\right)^{n_2}\frac{1}{n_3!}\left(\frac{1}{3!}\int J_1J_2J_3G^{(3)}_c\right)^{n_3}\cdots$$

such that $2n_2 + 3n_3 + \cdots = n$, in a hopefully obvious condensed notation. For example, $n = 4$ has $(n_2 = 2, n_{\neq 2} = 0) + (n_4 = 1, n_{\neq 4} = 0)$, and $n = 6$ has $(n_6 = 1, n_{\neq 6} = 0) + (n_4 = 1, n_2 = 1, n_{\neq 4,2} = 0) + (n_3 = 2, n_{\neq 3} = 0) + (n_2 = 3, n_{\neq 2} = 0)$. Consider the term with $G^{(k)}_c$: take

$$\frac{\delta^{kn_k}}{\delta J(x_1)\cdots\delta J(x_{kn_k})}\frac{1}{n_k!}\left(\frac{1}{k!}\int J_1\cdots J_kG^{(k)}_c\right)^{n_k}$$

This is completely symmetric under permutations of $x_1, \ldots, x_{kn_k}$. To make this explicit we rewrite it as

$$\frac{1}{n_k!}\left(\frac{1}{k!}\int J_1 \cdots J_k G_c^{(k)}\right)^{n_k} = \frac{1}{n_k!}\frac{1}{(k!)^{n_k}}$$
$$\times \int \prod_{i=1}^{kn_k} d^4y_i \, J(y_1)\cdots J(y_{kn_k}) G_c^{(k)}(y_1, \ldots, y_k) \cdots G_c^{(k)}(y_{(n_k-1)k+1}, \ldots, y_{n_k k})$$

Taking kn_k J-derivatives we obtain

$$\frac{1}{n_k!}\frac{1}{(k!)^{n_k}} G_c^{(k)}(x_1, \ldots, x_k) \cdots G_c^{(k)}(x_{(n_k-1)k+1}, \ldots, x_{n_k k}) + \text{permutations of } x_1, \ldots, x_{n_k k}$$

This contains many repeated terms. We need to count the number of inequivalent permutations. For each $G_c^{(k)}$ there are $k!$ equivalent permutations of the arguments; this gives $(k!)^{n_k}$. Then we can permute the $G_c^{(k)}$ among themselves; there are $n_k!$ such permutations. So we obtain

$$\frac{\delta^{kn_k}}{\delta J(x_1)\cdots\delta J(x_{kn_k})}\frac{1}{n_k!}\left(\frac{1}{k!}\int J_1 \cdots J_k G_c^{(k)}\right)^{n_k}$$
$$= G_c^{(k)}(x_1, \ldots, x_k) \cdots G_c^{(k)}(x_{(n_k-1)k+1}, \ldots, x_{n_k k}) + \text{inequiv-perms}$$

Finally combine all terms and symmetrize over $x_1, \ldots, x_n$. We obtain all possible combinations of G_c's that can make $G^{(n)}$. But that is precisely what we intended to show.

Chapter 6

Fields that are not scalars

6.1 Generalities

So far we have concentrated our studies on fields that transform very simply under Lorentz transformations. Brief review: we want $\phi(x)$ to correspond to $\phi'(x')$ when $x' = \Lambda x$. For a scalar field "correspond to" means they are equal, $\phi'(x') = \phi(x)$. Tha is,

$$\phi'(x) = \phi(\Lambda^{-1}x)\,.$$

Less trivial is the case of a vector field, $A^\mu(x)$. We can obtain it from taking a derivative on the scalar field which gives, $\partial_\mu \phi'(x) = \partial_\nu \phi(\Lambda^- 1x)(\Lambda^{-1})^\nu{}_\mu = \Lambda_\mu{}^\nu \partial_\nu \phi(\Lambda^- 1x)$. This holds for any vector so

$$A'^\mu(x') = \Lambda^\mu{}_\nu A^\nu(x) \qquad \text{or} \qquad A'^\mu(x) = \Lambda^\mu{}_\nu A^\nu(\Lambda^{-1}x)\,.$$

We can generalize this to other tensors easily, by considering tensor products of vectors, *e.g.*,

$$B'^{\mu\nu\lambda}(x) = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma \Lambda^\lambda{}_\delta B^{\rho\sigma\delta}(\Lambda^{-1}x)$$

In general, a collection of fields ψ_α, $\alpha = 1, \ldots, n$ transforms as

$$\psi'_\alpha(x) = D_{\alpha\beta}(\Lambda)\psi_\beta(\Lambda^{-1}x)\,, \tag{6.1}$$

where $D_{\alpha\beta}$ is an $n \times n$ matrix function of Λ. If $\psi'_\alpha(x') = D_{\alpha\beta}(\Lambda_1)\psi_\beta(x)$ with $x' = \Lambda_1 x$ and $\psi''_\alpha(x) = D_{\alpha\beta}(\Lambda_2)D_{\beta\gamma}(\Lambda_1)\psi_\gamma(\Lambda_1^{-1}\Lambda_2^{-1}x)$. This should equal the transformation with $x'' = \Lambda_2\Lambda_1 x$, $\psi''_\alpha(x) = D_{\alpha\beta}(\Lambda_2\Lambda_1)\psi_\beta((\Lambda_2\Lambda_1)^{-1}x)$. This imposes a requirement on the functions $D(\Lambda)$ that they furnish a *representation* of the Lorentz group:

$$D(\Lambda_2)D(\Lambda_1) = D(\Lambda_2\Lambda_1) \tag{6.2}$$

It suffices to understand the irreducible representations. Brief review/introduction. If D is a representation, so is SDS^{-1} for any invertible matrix S. If we can find an S such that SDS^{-1} is block diagonal for all Λ,

$$\begin{pmatrix} D^{(1)} & 0 & \cdots & 0 \\ 0 & D^{(2)} & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & D^{(N)} \end{pmatrix}$$

then we say D is *reducible* (to be more precise, the case that $D^{(1)}$ is the only block in the block diagonal matrix should be excluded). Else, it is *irreducible*. If D is reducible it (or, rather, SDS^{-1} for some S) can be written as the *direct sum* of irreducible representations $D^{(i)}$, $i = 1, \ldots, N$, and we write this as $D = D^{(1)} \oplus D^{(2)} \oplus \cdots \oplus D^{(N)}$. The point is that one can form any reducible representation from knowledge of the possible irreducible ones. So we need only determine the fields that correspond to irreducible representations,

$$\psi'^{(1)}(x) = D^{(1)}(\Lambda)\psi^{(1)}(\Lambda^{-1}x), \ldots, \psi'^{(N)}(x) = D^{(N)}(\Lambda)\psi^{(1)}(\Lambda^{-1}x)\,,$$

and the whole collection transforms as a reducible representation

$$\begin{pmatrix} \psi'^{(1)}(x) \\ \psi'^{(2)}(x) \\ \vdots \\ \psi'^{(N)}(x) \end{pmatrix} = \begin{pmatrix} D^{(1)} & 0 & \cdots & 0 \\ 0 & D^{(2)} & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & D^{(N)} \end{pmatrix} \begin{pmatrix} \psi^{(1)}(\Lambda^{-1}x) \\ \psi^{(2)}(\Lambda^{-1}x) \\ \vdots \\ \psi^{(N)}(\Lambda^{-1}x) \end{pmatrix}$$

An example: a pair of scalars, ϕ_1, ϕ_2, a vector, A^μ, and a tensor, $T^{\mu\nu}$, with the transformations given above.

If D acts on d dimensional vectors we say the dimension of D is d, $\dim(D) = d$.

(Aside: The representations may be double valued. You have seen this in QM. A spin-$\frac{1}{2}$ wave-function, $\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$ transforms, under rotations by an angle θ about the $\hat{n}$ axis, by $\psi \to e^{i\frac{1}{2}\theta\hat{n}\cdot\vec{\sigma}}\psi$. That is $D(R) = e^{i\frac{1}{2}\theta\hat{n}\cdot\vec{\sigma}}$ is a 2-dim representation of the rotation R by an angle θ about the $\hat{n}$ axis. But R is the same for $\theta = 0$ and $\theta = 2\pi$, and $e^{i\pi\hat{n}\cdot\vec{\sigma}} = -1$.)

If $D(\Lambda)$ is a representation, so is $D^*(\Lambda)$. Proof: take the complex conjugate of $D(\Lambda_1)D(\Lambda_2) = D(\Lambda_1\Lambda_2)$. Clearly, $\dim(D^*) = \dim(D)$. D^* may or may not be equivalent to D.

If $D^{(1)}$ and $D^{(2)}$ are representations, so is the *tensor product*, $D^{(1)} \otimes D^{(2)}$. The tensor product is defined as always: if $D^{(1)}$ acts on $\psi^{(1)}$, $D^{(2)}$ on $\psi^{(2)}$, then $D^{(1)} \otimes D^{(2)}$ acts on $\psi^{(1)} \otimes \psi^{(2)}$ according to $(D^{(1)} \otimes D^{(2)})(\psi^{(1)} \otimes \psi^{(2)}) = (D^{(1)}\psi^{(1)}) \otimes$

$(D^{(2)}\psi^{(2)})$. If $d = \dim(D)$ and $d' = \dim(D')$ then $\dim(D \oplus D') = d + d'$ and $\dim(D\otimes D') = dd'$. Generally, $D\otimes D'$ is reducible, $D\otimes D' = D^{(1)}\oplus D^{(2)}\oplus\cdots\oplus D^{(N)}$, with $d_1 + d_2 + \cdots + d_N = dd'$. An example is a two index tensor. It can be obtained from the tensor product of two vectors, $C^{\mu\nu} = A^\mu B^\nu$, and is a reducible representation: under Lorentz transformations the trace $\eta_{\mu\nu}C^{\mu\nu}$, the anti-symmetric part, $C^{[\mu\nu]} = \frac{1}{2}(C^{\mu\nu} - C^{\nu\mu})$, and the symmetric traceless part, $C^{\{\mu\nu\}} - \frac{1}{4}\eta^{\mu\nu}\eta_{\lambda\sigma}C^{\lambda\sigma} = \frac{1}{2}(C^{\mu\nu} + C^{\nu\mu}) - \frac{1}{4}\eta^{\mu\nu}\eta_{\lambda\sigma}C^{\lambda\sigma}$ do not mix into each other. Moreover, one can show that the 6-dimensional antisymmetric 2-index tensor is itself the direct sum of two 3-dimensional irreducible representations, satisfying $C^{\mu\nu}_{\pm} = \pm\frac{1}{2}\epsilon^{\mu\nu}{}_{\lambda\sigma}C^{\lambda\sigma}_{\pm}$ (these are said to be self-dual and anti-self-dual, respectively). The $4 \times 4 = 16$-dimensional tensor product of two 4-vectors splits into four irreducible representations of dimension 1, 3, 3, and 9, as we have just seen. With a slight abuse of notation, this is $4 \otimes 4 = 1 \oplus 3 \oplus 3' \oplus 9$.

Here is the point. We can build up every irreducible representation (and from them every reducible representation, hence every representation) by starting from some small basic representations, and taking their tensor products repeatedly: these tensor products are direct sums of new irreducible representation, and the more basic representations we tensor-product the higher the dimension of the new irreducible representations we will find.

It turns out, as we will show below, the representations of the Lorentz groups are labeled by two half-integers, (s_+, s_-) and have dimension $(2s_+ + 1)(2s_- + 1)$. This is because the Lorentz group (or rather, its algebra) is isomorphic to two copies of spin, $SU(2)\times SU(2)$, and as you know from particle QM the representations of spin are classified by half integer $s = 0, \frac{1}{2}, 1, \ldots$ and have dimension $2s + 1 = 1, 2, 3, \ldots$ correspondingly. For example, $(0, 0)$ is a 1-dimensional representation, the scalar, $(\frac{1}{2}, \frac{1}{2})$ is a 4-dimensional representation, corresponding to vectors, A^μ. For the tensor product of two vectors, we need first, from QM, that $\frac{1}{2} \otimes \frac{1}{2} = 0 \oplus 1$. So $(\frac{1}{2}, \frac{1}{2})\otimes(\frac{1}{2}, \frac{1}{2}) = (0, 0)\oplus(1, 0)\oplus(0, 1)\oplus(1, 1)$. The two 3-dimensional representations associated with the antisymmetric tensor correspond to $(1, 0)$ and $(0, 1)$ and the 9-dimensional symmetric traceless tensor is $(1, 1)$.

6.2 Spinors

Back to Lorentz group and physics. Question: other than scalars and tensors, what other representations of the Lorentz group do we have? Answer: spinors and their tensor products. Let $\sigma^\mu = (\sigma^0, \sigma^i)$, where

$$\sigma^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Let
$$P = p_\mu \sigma^\mu = \begin{pmatrix} p^0 - p^3 & -p^1 + ip^2 \\ -p^1 - ip^2 & p^0 + p^3 \end{pmatrix}.$$
Then $\det(P) = p^2 = \eta_{\mu\nu} p^\mu p^\nu$. Then for any unimodular 2×2 matrix $\widetilde{M}$ we have $\det(\widetilde{M}^\dagger P \widetilde{M}) = \det(P)$. Since any hermitian matrix can be expanded in terms of σ^μ with real coefficients and $P' = \widetilde{M}^\dagger P \widetilde{M}$ is hermitian (if P is), then $P' = p'_\mu \sigma^\mu$ with $p'^2 = p^2$. That is, $\widetilde{M} = \widetilde{M}(\Lambda)$ induces a Lorentz transformation.

Now $\widetilde{M}(\Lambda)$ does not satisfy the representation defining equation (6.2). If $p'' = \Lambda_2 p'$ and $p' = \Lambda_1 p$ then $P'' = \widetilde{M}^\dagger(\Lambda_2) P' \widetilde{M}(\Lambda_2) = \widetilde{M}^\dagger(\Lambda_2)\widetilde{M}^\dagger(\Lambda_1) P \widetilde{M}(\Lambda_1)\widetilde{M}(\Lambda_2)$ and this should correspond to $p'' = \Lambda_2\Lambda_1 p$ or $P'' = \widetilde{M}^\dagger(\Lambda_2\Lambda_1) P \widetilde{M}(\Lambda_2\Lambda_1)$. Comparing we see that,
$$\widetilde{M}(\Lambda_1)\widetilde{M}(\Lambda_2) = \widetilde{M}(\Lambda_2\Lambda_1)\,.$$
Defining $M(\Lambda) = \widetilde{M}(\Lambda^{-1})$ we obtain a representation,
$$M(\Lambda_2)M(\Lambda_1) = \widetilde{M}(\Lambda_2^{-1})\widetilde{M}(\Lambda_1^{-1}) = \widetilde{M}(\Lambda_1^{-1}\Lambda_2^{-1}) = \widetilde{M}((\Lambda_2\Lambda_1)^{-1}) = M(\Lambda_2\Lambda_1)\,.$$

So $M(\Lambda)$ gives a 2-dimensional representation of Λ. It acts on 2-dimensional *spinors* (vectors in a 2-dimensional space):
$$\psi'_\alpha = M_{\alpha\beta}(\Lambda)\psi_\beta$$
If $P' = \widetilde{M}^\dagger P \widetilde{M}$ with $P' = p'_\nu \sigma^\nu = \Lambda_\nu{}^\mu p_\mu \sigma^\nu$ then equating this to $M^\dagger P M$ for arbitrary p_μ we must have
$$\widetilde{M}(\Lambda)^\dagger \sigma^\mu \widetilde{M}(\Lambda) = \Lambda_\nu{}^\mu \sigma^\nu = (\Lambda^{-1})^\mu{}_\nu \sigma^\nu \tag{6.3}$$
or
$$M(\Lambda)^\dagger \sigma^\mu M(\Lambda) = \Lambda^\mu{}_\nu \sigma^\nu \tag{6.4}$$

As we have seen, the complex conjugate, $M^*(\Lambda)$, must also be a representation. It satisfies,
$$M^T \sigma^{\mu *} M^* = \Lambda^\mu{}_\nu \sigma^{\nu *}$$
or sandwiching with σ^2 —a similarity transformation— and defining $\bar\sigma^\mu = \sigma^2\sigma^{\mu *}\sigma^2 = (\sigma^0, -\sigma^i)$, and $\overline{M} = \sigma^2 M^* \sigma^2$:
$$\overline{M}^\dagger \bar\sigma^\mu \overline{M} = \Lambda^\mu{}_\nu \bar\sigma^\nu$$

So there must be a 2-dimensional representation of the Lorentz group, of 2-component vectors, or *spinors*, that transform according to
$$\psi'_\alpha(x) = M_{\alpha\beta}(\Lambda)\psi_\beta(\Lambda^{-1}x)$$

and also a complex conjugate representation that acts on other 2-component vectors, also called a spinors, that transform according to

$$\overline{\chi}'_\alpha(x) = \overline{M}_{\alpha\beta}(\Lambda)\overline{\chi}_\beta(\Lambda^{-1}x)\,.$$

It is convenient to arrange the spinors into 2-component column vectors, and write

$$\psi' = M(\Lambda)\psi \quad \text{and} \quad \overline{\chi}' = \overline{M}(\Lambda)\overline{\chi}$$

Note that

$$\psi'^\dagger\sigma^\mu\psi' = \psi^\dagger M^\dagger\sigma^\mu M\psi = \Lambda^\mu{}_\nu\,\psi^\dagger\sigma^\nu\psi\,.$$

and

$$\overline{\chi}'^\dagger\overline{\sigma}^\mu\overline{\chi}' = \overline{\chi}^\dagger\overline{M}^\dagger\overline{\sigma}^\mu\overline{M}\overline{\chi} = \Lambda^\mu{}_\nu\,\overline{\chi}^\dagger\overline{\sigma}^\nu\overline{\chi}\,.$$

In making tensors out of these it is convenient to distinguish the indices in ψ and $\overline{\chi}$, so we write ψ_α and $\overline{\chi}_{\dot\alpha}$ (still with $\alpha = 1,2$ and $\dot\alpha = 1,2$).

We can construct a scalar out of two spinors, ψ_1 and ψ_2:

$$\psi_1'^T\sigma^2\psi_2' = \psi_1^T M^T\sigma^2 M\psi_2 = \psi_1^T\sigma^2\psi_2 \tag{6.5}$$

The last step follows from

$$\begin{aligned}(M^T\sigma^2 M)_{\alpha\beta} &= -i\epsilon_{\gamma\delta}M_{\alpha\gamma}M_{\beta\delta}\\ &= -i\epsilon_{\alpha\beta}\det(M)\\ &= \sigma^2_{\alpha\beta}\end{aligned}$$

Similarly, $\overline{\chi}_1^T\sigma^2\overline{\chi}_2$ is a scalar. Note that $\psi_1^\dagger\psi_2$ is not a scalar since M is not generally unitary. Note also that $\psi^T\sigma^2\psi = 0$, so we cannot make a scalar out of a single ψ.

Form Eq. (6.4) we can verify that

$$M = \exp\left(-\tfrac{i}{2}\vec\alpha\cdot\vec\sigma\right) = \cos(\tfrac12\alpha) - i\hat\alpha\cdot\vec\sigma\,\sin(\tfrac12\alpha) \tag{6.6}$$

is a representation of $\Lambda = R =$ a rotation by angle α about an axis in the $\hat\alpha = \vec\alpha/\alpha$ direction, and

$$M = \exp\left(\tfrac12\vec\beta\cdot\vec\sigma\right) = \cosh(\tfrac12\beta) + \hat\beta\cdot\vec\sigma\,\sinh(\tfrac12\beta)$$

is a representation of a boost by velocity β, that is, a representation of

$$\Lambda = \begin{pmatrix}\gamma & \gamma\beta\\ \gamma\beta & \gamma\end{pmatrix}$$

where $\gamma = 1/\sqrt{1-\beta^2}$, $\vec\beta$ is along the x axis, and the last two rows and columns of Λ have been omitted. By complex conjugating it follows that

$$\overline{M} = \exp\left(-\tfrac{i}{2}\vec\alpha\cdot\vec\sigma\right) \tag{6.7}$$

is a representation of the same rotation R (a rotation by angle α about $\hat{\alpha}$), and

$$\overline{M} = \exp\left(-\tfrac{1}{2}\vec{\beta}\cdot\vec{\sigma}\right)$$

is a representation of a boost by velocity $\vec{\beta}$.

6.3 A Lagrangian for Spinors

We want to construct a Lagrangian density for spinors. We put the general constraints:

(i) Constructed from ψ and $\psi^\dagger$ and their first derivatives $\partial_\mu\psi$ and $\partial_\mu\psi^\dagger$.

(ii) Real, $\mathcal{L}^* = \mathcal{L}$ (or for quantum fields, hermitian, $\mathcal{L}^\dagger = \mathcal{L}$).

(iii) Lorentz invariant (at least up to total derivatives)

(iv) At most quadratic in the fields.

The last condition is not generally necessary. We impose it for simplicity. Higher orders in the fields will correspond to interactions.

From (iv) we need to construct $\mathcal{L}$ from bilinears $\psi^\dagger\otimes\psi$ and $\psi\otimes\psi$. We know we can form a vector out of these, but not a scalar. Now,

$$\partial_\mu\psi(\Lambda^{-1}x) = \partial_\mu((\Lambda^{-1})^\nu{}_\lambda x^\lambda)(\partial_\nu\psi)(\Lambda^{-1}x) = (\Lambda^{-1})^\nu{}_\mu(\partial_\nu\psi)(\Lambda^{-1}x)\,,$$

so that

$$\psi'^\dagger\sigma^\mu\partial_\mu\psi' = \psi^\dagger M^\dagger\sigma^\mu M(\Lambda^{-1})^\nu{}_\mu\partial_\nu\psi = \psi^\dagger\Lambda^\mu{}_\lambda\sigma^\lambda M(\Lambda^{-1})^\nu{}_\mu\partial_\nu\psi = \psi^\dagger\sigma^\mu\partial_\mu\psi\,.$$

Both $\psi^\dagger\sigma^\mu\partial_\mu\psi$ and $\partial_\mu\psi^\dagger\sigma^\mu\psi$ transform as scalars so they are candidates for an invariant Lagrangian. But the sum is a total derivative, $\partial_\mu(\psi^\dagger\sigma^\mu\psi)$ so it is irrelevant (it does not contribute to the equations of motion). So we take the difference,

$$\tfrac{1}{2}\psi^\dagger\sigma^\mu\overleftrightarrow{\partial_\mu}\psi = \tfrac{1}{2}\left(\psi^\dagger\sigma^\mu\partial_\mu\psi - \partial_\mu\psi^\dagger\sigma^\mu\psi\right)$$

as a possible term in the Lagrangian. It's complex conjugate is

$$\begin{aligned}\tfrac{1}{2}(\psi^\dagger\sigma^\mu\overleftrightarrow{\partial_\mu}\psi)^* &= \tfrac{1}{2}\left(\psi^T\sigma^{\mu*}\partial_\mu\psi - \partial_\mu\psi^T\sigma^{\mu*}\psi^*\right)\\ &= \tfrac{1}{2}\left(\partial_\mu\psi^\dagger\sigma^{\mu\dagger}\psi - \psi^\dagger\sigma^{\mu\dagger}\partial_\mu\psi\right)\\ &= -\tfrac{1}{2}\psi^\dagger\sigma^\mu\overleftrightarrow{\partial_\mu}\psi\end{aligned}$$

If we take $\mathcal{L} = A\frac{1}{2}\psi^\dagger\sigma^\mu \overleftrightarrow{\partial_\mu}\psi$ then for $\mathcal{L}^* = \mathcal{L}$ we must have $A = \pm i|A|$. Redefining $\psi \to \frac{1}{\sqrt{|A|}}\psi$ we have a candidate Lagrangian density

$$\mathcal{L} = \pm\frac{1}{2}\psi^\dagger\sigma^\mu\, i\overleftrightarrow{\partial_\mu}\psi$$

We will see below how to choose properly between the two signs. Note that if we relax assumption (iv) we could add other terms, *e.g.,* $(\psi^\dagger\sigma^\mu\psi)(\psi^\dagger\sigma_\mu\psi)$.

Equations of motion: recall, with a complex field we can take variations with respect to ψ and $\psi^\dagger$ separately, as if they were independent variables. Now in the action integral it is convenient to integrate by parts, so that

$$\int d^4x\,\mathcal{L} = \pm\int d^4x\,\frac{1}{2}\psi^\dagger\sigma^\mu\, i\overleftrightarrow{\partial_\mu}\psi = \pm\int d^4x\,\psi^\dagger\sigma^\mu\, i\partial_\mu\psi$$

Then, trivially,

$$\frac{\delta\mathcal{L}}{\delta\psi^\dagger} = 0 \quad\Rightarrow\quad \sigma^\mu\partial_\mu\psi = 0\,.$$

This is

$$(\sigma^0\partial_0 + \sigma^i\partial_i)\psi = 0 \quad\Rightarrow\quad \partial_0\psi = -\sigma^i\partial_i\psi = -\vec{\sigma}\cdot\vec{\nabla}\psi$$

so that

$$\partial_0^2\psi = -\partial_0\vec{\sigma}\cdot\vec{\nabla}\psi = -\vec{\sigma}\cdot\vec{\nabla}\partial_0\psi = (\vec{\sigma}\cdot\vec{\nabla})^2\psi = \nabla^2\psi$$

This is precisely the KG equation with $m = 0$

$$\partial^2\psi = (\partial_0^2 - \nabla^2)\psi = 0\,.$$

Each component of ψ satisfies the massless KG equation.

We can construct a Lagrangian for fields in the complex conjugate representation, $\bar{\chi}$. An analogous argument gives us

$$\mathcal{L} = \pm\bar{\chi}^\dagger\bar{\sigma}^\mu\, i\partial_\mu\bar{\chi}$$

with equation of motion

$$\bar{\sigma}^\mu\partial_\mu\bar{\chi} = 0$$

This again gives $\partial^2\bar{\chi} = 0$ but now with $\partial_0\bar{\chi} = \vec{\sigma}\cdot\vec{\nabla}\bar{\chi}$ (note the sign difference).

Plane wave expansion. Since these are complex fields we have different coefficients for the positive and negative energy components:

$$\psi_\alpha = \int(dk)\left[e^{-ik\cdot x}B_{\vec{k},\alpha} + e^{ik\cdot x}D_{\vec{k},\alpha}\right]$$

where $B_{\vec{k},\alpha}$ and $D_{\vec{k},\alpha}$ are two component operator valued objects. Since $\partial^2\psi = 0$ we must have $k^2 = E_{\vec{k}}^2 - \vec{k}^2 = 0$, that is, $k^0 = E_{\vec{k}} = |\vec{k}|$ as we should for massless particles. But the equation of motion is first order in derivatives. Using $\partial_\mu e^{\mp ik\cdot x} = \mp ik_\mu e^{\mp ik\cdot x}$ we have

$$\int (dk)\left[e^{-ik\cdot x}k_\mu\sigma^\mu B_{\vec{k}} - e^{i\cdot x}k_\mu\sigma^\mu D_{\vec{k}}\right] = 0$$

That is, we need

$$\begin{pmatrix} k^0 - k^3 & -k^1 + ik^2 \\ -k^1 - ik^2 & k^0 - k^3 \end{pmatrix}\begin{pmatrix} B_{\vec{k},1} \\ B_{\vec{k},2}\end{pmatrix} = 0 \qquad \begin{array}{c}(k^0 - k^3)B_{\vec{k},1} = -(-k^1 + ik^2)B_{\vec{k},2} \\ (-k^1 - ik^2)B_{\vec{k},1} = -(k^0 + k^3)B_{\vec{k},2}\end{array}$$

and similarly for $D_{\vec{k},\alpha}$. Let

$$u_{\vec{k}} = N_{\vec{k}}\begin{pmatrix} k^1 - ik^2 \\ k^0 - k^3\end{pmatrix}.$$

$N_{\vec{k}}$ is a normalization factor. We could choose it to have, for example, $u_{\vec{k}}^\dagger u_{\vec{k}} = 1$, but we will wait to choose it conveniently later. Note that since $k^2 = 0$ we can also write this as

$$u_{\vec{k}} = N_{\vec{k}}\frac{k^0 - k^3}{k^1 - ik^2}\begin{pmatrix} k^0 + k^3 \\ k^1 + ik^2\end{pmatrix} = N'_{\vec{k}}\begin{pmatrix} k^0 + k^3 \\ k^1 + ik^2\end{pmatrix}.$$

For the plane wave-expansion of $\bar{\chi}$ we need to solve $k_\mu\bar{\sigma}^\mu v_{\vec{k}} = 0$. But this is just like the equation for $u_{\vec{k}}$ only replacing $-\vec{k}$ for $\vec{k}$. So $v_{\vec{k}} = u_{-\vec{k}}$ up to a phase. Since there is only one solution to the matrix equation, we rewrite our plane wave expansion as

$$\psi = \int (dk)\left[e^{-ik\cdot x}\beta_{\vec{k}}u_{\vec{k}} + e^{ik\cdot x}\delta_{\vec{k}}^\dagger u_{\vec{k}}\right]$$

where now the operators $\beta_{\vec{k}}$ and $\delta_{\vec{k}}$ are one-component objects. So other than the spinor coefficient $u_{\vec{k}}$ this expansion looks very much like the one for complex scalar fields.

6.3.1 Hamiltonian; Fermi-Dirac Statistics

Let's compute the Hamiltonian. This should allow us to fix the sign in the Lagrangian density since we want a Hamiltonian that is bounded from below. From the density

$$\mathcal{H} = \frac{\partial\mathcal{L}}{\partial(\partial_t\psi)}\partial_t\psi - \mathcal{L} = \mp i\psi^\dagger\vec{\sigma}\cdot\vec{\nabla}\psi$$

we obtain the Hamiltonian in terms of creation/annihilation operators,

$$H = \mp \int (dk)(dk') u^\dagger_{\vec{k}'} \vec{\sigma}\cdot\vec{k}\, u_{\vec{k}} \left[(2\pi)^3\delta(\vec{k}'-\vec{k})(\delta_{\vec{k}'}\delta^\dagger_{\vec{k}} - \beta^\dagger_{\vec{k}'}\beta_{\vec{k}}) + (2\pi)^3\delta(\vec{k}'+\vec{k})(\beta^\dagger_{\vec{k}'}\delta^\dagger_{\vec{k}} - \delta_{\vec{k}'}\beta_{\vec{k}})\right]$$

Now use, $k_\mu\sigma^\mu u_{\vec{k}} = 0$ or

$$\vec{k}\cdot\vec{\sigma}\, u_{\vec{k}} = E_{\vec{k}} u_{\vec{k}} \qquad \text{with } E_{\vec{k}} = |\vec{k}|,$$

and adopt the normalization

$$u^\dagger_{\vec{k}} u_{\vec{k}} = 2E_{\vec{k}}\,.$$

Moreover, from the explicit form of $u_{\vec{k}}$ we have

$$u^\dagger_{\vec{k}} u_{-\vec{k}} = 0$$

Putting these together we have

$$H = \mp \int (dk) E_{\vec{k}} \left(-\beta^\dagger_{\vec{k}}\beta_{\vec{k}} + \delta_{\vec{k}}\delta^\dagger_{\vec{k}}\right)$$

If we take, as we have done before,

$$[\beta_{\vec{k}}, \beta^\dagger_{\vec{k}'}] = (2\pi)^3 2E_{\vec{k}}\delta^{(3)}(\vec{k}-\vec{k}') = [\delta_{\vec{k}}, \delta^\dagger_{\vec{k}'}] \tag{6.8}$$

then

$$H = \mp \int (dk) E_{\vec{k}} \left(-\beta^\dagger_{\vec{k}}\beta_{\vec{k}} + \delta^\dagger_{\vec{k}}\delta_{\vec{k}}\right)$$

plus an infinite constant that we throw away (normal ordering). We were hoping to fix the sign, but instead we encounter a disaster! If we choose the $-$ sign in the definition of $\mathcal{L}$ the $\beta^\dagger_{\vec{k}}\beta_{\vec{k}}$ has a spectrum unbounded from below while if we choose the $+$ sign the $\delta^\dagger_{\vec{k}}\delta_{\vec{k}}$ term is unbounded form below.

But if instead of (6.8) we choose anti-commutation relations,

$$\{\beta_{\vec{k}}, \beta^\dagger_{\vec{k}'}\} = (2\pi)^3 2E_{\vec{k}}\delta^{(3)}(\vec{k}-\vec{k}') = \{\delta_{\vec{k}}, \delta^\dagger_{\vec{k}'}\} \tag{6.9}$$

where $\{A, B\} \equiv AB + BA$, then up to an infinite constant

$$H = \pm \int (dk) E_{\vec{k}} \left(\beta^\dagger_{\vec{k}}\beta_{\vec{k}} + \delta^\dagger_{\vec{k}}\delta_{\vec{k}}\right)$$

This is bounded from below only if we take the $+$ sign in $\mathcal{L}$, unbounded form below otherwise. This fixes the sign,

$$\mathcal{L} = \psi^\dagger\, i\sigma^\mu\partial_\mu\psi\,,$$

and furthermore tells us that the fields anti-commute,

$$\{\psi^\dagger(x), \psi(y)\}|_{x^0=y^0} = \delta^{(3)}(\vec{x}-\vec{y})$$

The Fock space consists of particles $\beta^\dagger_{\vec{k}}|0\rangle$ and antiparticles, $\delta^\dagger_{\vec{k}}|0\rangle$. Two particle states satisfy $|\vec{k}_1, \vec{k}_2\rangle = \beta^\dagger_{\vec{k}_1}\beta^\dagger_{\vec{k}_2}|0\rangle = -\beta^\dagger_{\vec{k}_2}\beta^\dagger_{\vec{k}_1}|0\rangle = -|\vec{k}_2, \vec{k}_1\rangle$, so the wave function is anti-symmetric; so are $\beta^\dagger_{\vec{k}_1}\delta^\dagger_{\vec{k}_2}|0\rangle$ and $\delta^\dagger_{\vec{k}_1}\delta^\dagger_{\vec{k}_2}|0\rangle$. We have discovered Fermi-Dirac statistics! Moreover, we were forced into it by consistency of the theory. This spin-statistics connection is a theorem in QFT, rather than an *ad-hoc* rule as in particle QM.

A note on the normalization of $u_{\vec{k}}$. It seems that we chose it as $2E_{\vec{k}}$ to give the Hamiltonian as the sum of $E_{\vec{k}}$ times the number of modes. Actually, we should fix the normalization to give $\{\psi^\dagger(x), \psi(y)\}|_{x^0=y^0} = i\delta^{(3)}(\vec{x}-\vec{y})$ given that creation/annihilation operators satisfy (6.9). It is a simple exercise to check that this is the case:

$$\{\psi^\dagger_\alpha(x), \psi_\beta(y)\}|_{x^0=y^0} = \int (dk) e^{i\vec{k}\cdot\vec{k}} \left[u^*_{\vec{k}\alpha}u_{\vec{k}\beta} + v^*_{\vec{k}\alpha}v_{\vec{k}\beta}\right]$$

where we have used $v_{\vec{k}} = u_{-\vec{k}}$. We can now check by direct computation that

$$u_{\vec{k}}u^\dagger_{\vec{k}} + v_{\vec{k}}v^\dagger_{\vec{k}} = 2E_{\vec{k}}\mathbb{1} \tag{6.10}$$

giving the desired result. The relation (6.10) is in fact a completeness relation,

$$\sum \frac{|n\rangle\langle n|}{\langle n|n\rangle} = \mathbb{1} \qquad \text{is} \qquad \frac{u_{\vec{k}}u^\dagger_{\vec{k}} + v_{\vec{k}}v^\dagger_{\vec{k}}}{2E_{\vec{k}}} = \mathbb{1}$$

Helicity(η).

Helicity of a state is defined as the angular momentum along the direction of motion. Take a particle moving along the z-axis, $k^1 = k^2 = 0$, $k^3 = \pm k^0$. Then if the particle moves in the z direction, $k^3 = k^0$

$$u^{(+)} \propto \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

while if it moves in the negative z-direction, $k^3 = -k^0$,

$$u^{(-)} \propto \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

For a rotation by α about the z-axis, Eq. (6.6) gives

$$M\begin{pmatrix}1\\0\end{pmatrix} = e^{-i\alpha/2}\begin{pmatrix}1\\0\end{pmatrix} \qquad \text{and} \qquad M\begin{pmatrix}0\\1\end{pmatrix} = e^{i\alpha/2}\begin{pmatrix}0\\1\end{pmatrix},$$

so that $u^{(\pm)}$ are eigenvectors of M with eigenvalues $e^{\mp i\alpha/2}$. But

$$M = e^{-i\alpha J^3}$$

so $u^{(+)}$ is an eigenvector of $\eta = \hat{k}\cdot\vec{J} = J^3$ with eigenvalue, or *helicity*, $\eta = \frac{1}{2}$ while $u^{(-)}$ has $\eta = \hat{k}\cdot J = -J^3$ so again $\eta = \frac{1}{2}$.

(Aside: In general

$$J^{\mu\nu} = L^{\mu\nu} + S^{\mu\nu} = \int d^3x\, [x^\mu T^{0\nu} - x^\nu T^{0\mu} - i\pi\mathcal{J}^{\mu\nu}\phi]$$

where $\phi'(x) = (1-\frac{i}{2}\omega_{\mu\nu}\mathcal{J}^{\mu\nu})\phi(x-\omega x)$ and π is the canonical momentum conjugate to ϕ. For the complex case one must sum over π and $\pi^\dagger$. For spinors we only have $\pi = \partial\mathcal{L}/\partial(\partial_t\psi) = \psi^\dagger\, i\sigma^0 = i\psi^\dagger$, since $\pi^\dagger = \partial\mathcal{L}/\partial(\partial_t\psi^\dagger) = 0$. So for spinors, $S^{\mu\nu} = \int d^3x\, \psi^\dagger\mathcal{M}^{\mu\nu}\psi$. If $U = \exp(-\frac{i}{2}\omega_{\mu\nu}\mathcal{J}^{\mu\nu})$, then infinitesimally $U\phi(x)U^\dagger = (1-\frac{i}{2}\omega_{\mu\nu}\mathcal{J}^{\mu\nu})\phi(x-\omega x)$. For a rotation about the z-axis by angle θ take $\omega_{12} = -\omega_{21} = \theta$, so that $U = \exp(-i\theta J^3)$. This can be verified from $UA^\lambda(0)U^\dagger = (\eta^{\lambda\sigma} + \omega^{\lambda\sigma})A_\sigma(0)$:

$$\begin{aligned} A'^1 &= A^1 + \omega^{12}A_2 = A^1 - \omega^{12}A^2 = A^1 - \theta A^2 \\ A'^2 &= A^2 + \omega^{21}A_1 = A^2 - \omega^{21}A^1 = A^2 + \theta A^1 . \end{aligned}$$

The factor of 2 in the definition of the spin part, $S^{\mu\nu}$, is correct. This can be checked with the vector representation, for which $(\mathcal{J}^{\mu\nu})_{\lambda\sigma} = -i(\delta^\mu_\lambda\delta^\nu_\sigma - \delta^\nu_\lambda\delta^\mu_\sigma)$. End aside).

To determine the helicity of 1-particle states annihilated and created by ψ, it is convenient to project these states into or out of the vacuum. Consider then a transformation $U(R)$ by a rotation R by angle α about the z-directions and a state with momentum $\vec{k} = k\hat{z}$:

$$\langle 0|U(R)\psi(0)U^\dagger(R)|k\hat{z}\rangle = \langle 0|M(R)\psi(0)|k\hat{z}\rangle$$

The right hand side picks up only the contribution in the plane wave expansion from the annihilation operator with momentum $\vec{k} = k\hat{z}$, so the matrix M acts on this spinor giving a factor of $\exp(-i\alpha/2)$. On the left hand side we have $\langle 0|\psi(0)(U^\dagger(R)|k\hat{z}\rangle)$. Since R does not change $\vec{k} = k\hat{z}$, the state can change at most by an overall phase. Comparing we read off the phase, $U^\dagger(R)|k\hat{z}\rangle = \exp(-i\alpha/2)|k\hat{z}\rangle$ or $U(R)|k\hat{z}\rangle = \exp(i\alpha/2)|k\hat{z}\rangle$. But $U(R) = \exp(-i\alpha J^3) = \exp(-i\alpha\eta)$, so the state annihilated by ψ has $\eta = -\frac{1}{2}$. Similarly, $\langle k\hat{z}|U(R)\psi(0)U^\dagger(R)|0\rangle = \langle k\hat{z}|M\psi(0)|0\rangle = \exp(-i\alpha/2)\langle k\hat{z}|\psi(0)|0\rangle$, gives $U(R)|k\hat{z}\rangle = \exp(-i\alpha/2)|k\hat{z}\rangle$. So ψ creates states with $\eta = \frac{1}{2}$.

For $\bar{\chi}$ we use the expansion basis of spinors $v_{\vec{k}}$. Since these are obtained from $u_{\vec{k}}$ by $\vec{k} \to -\vec{k}$, we expect they have $\eta = -\frac{1}{2}$. It is trivial to verify this: from Eq. (6.7) for a rotation $\overline{M} = M$ so the computation is as above, except now for $\vec{k}$ along the$z$-axis we get $v^{(+)} \propto \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and similarly for $\vec{k}$ pointing in the negative z-direction we have $v^{(-)} \propto \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, both of which give positive helicity.

To summarize: *ψ annihilates states with $\eta = -\frac{1}{2}$ and creates states with $\eta = \frac{1}{2}$; $\bar{\chi}$ annihilates states with $\eta = \frac{1}{2}$ and creates states with $\eta = -\frac{1}{2}$.*

6.3.2 Weyl vs Majorana

We stated earlier (see Eq. (6.5) and comment below) that the Lagrangian for spinors does not admit a mass term. $(\sigma^2)^{\alpha\beta}\psi_\alpha\psi_\beta = 0$ because the antisymmetric matrix $(i\sigma^2)^{\alpha\beta} = \epsilon^{\alpha\beta}$ is traced with the symmetric one $\psi_\alpha\psi_\beta$. But now that we have discovered that consistent quantization requires that spinors anti-commute we must revise this assertion: for anti-commuting fields $\psi_\alpha\psi_\beta$ is anti-symmetric!

Consider then

$$\mathcal{L} = \psi^\dagger\, i\sigma^\mu\partial_\mu\psi - (m\psi^T\epsilon\psi + \text{h.c.}) \tag{6.11}$$

where $\epsilon = i\sigma^2$. Comments:

(i) Two component massless spinors are called Weyl spinors. Massive ones are called Majorana spinors.

(ii) The Lagrangian for the Majorana spinor, Eq. (6.11), has no $U(1)$ symmetry, $\psi \to e^{i\alpha}\psi$. The Weyl case does; additional interactions with other fields may or may not respect the symmetry. For example, if we have also a complex scalar ϕ an interaction term in the Lagrangian $g\phi\psi^T\epsilon\psi + g^*\phi^*\psi^\dagger\epsilon\psi^*$ respects the symmetry, but if instead we have a real scalar, then $g\phi\psi^T\epsilon\psi + g^*\phi\psi^\dagger\epsilon\psi^*$ does not. Neither Weyl nor Majorana spinors can describe the electron: one because it does not have a mass the other because it does not carry charge.

(iii) The phase of the mass term is completely arbitrary. Since the kinetic term is invariant under $\psi \to e^{i\alpha}\psi$ we can always make a redefinition of the field ψ that changes the coefficient of the mass term by a phase $m \to e^{2i\alpha}m$: we are free to choose m real and positive.

(iv) Helicity .... To be filled in later. But main points (a) same computation as before, but (b) frame dependent since massive particle can be boosted to reverse direction of motion without changing spin

6.4 The Dirac Field

We need a description of massive spinors that carry charge, like the electron. We accomplish this by a using two spinors, ψ and $\bar{\chi}$, both with the same charge, that is, both transforming the same way under a $U(1)$ transformation, $\psi \to e^{i\alpha}\psi$ and $\bar{\chi} \to e^{i\alpha}\bar{\chi}$. Then under a Lorentz transformation

$$\bar{\chi}^\dagger\psi \to \bar{\chi}^\dagger\overline{M}^\dagger M\psi = \bar{\chi}^\dagger\psi$$

The last step follows from

$$(\overline{M}^\dagger M)_{\alpha\beta} = -\epsilon_{\alpha\gamma}M_{\delta\gamma}\epsilon_{\delta\sigma}M_{\sigma\beta}$$

and since $\det(M) = 1$, $\epsilon_{\delta\sigma}M_{\delta\gamma}M_{\sigma\beta} = \epsilon_{\gamma\beta}$ so we are left with $-\epsilon_{\alpha\gamma}\epsilon_{\gamma\beta} = \delta_{\alpha\beta}$. So $\bar{c}hi^\dagger\psi$ is both Lorentz and $U(1)$ invariant.

Hence we take

$$\mathcal{L} = \psi^\dagger\sigma^\mu\, i\partial_\mu\psi + \bar{\chi}^\dagger\bar{\sigma}^\mu\, i\partial_\mu\bar{\chi} - m(\bar{\chi}^\dagger\psi + \psi^\dagger\bar{\chi})$$

Equations of Motion:

$$\sigma^\mu\, i\partial_\mu\psi = m\bar{\chi} \qquad \text{and} \qquad \bar{\sigma}^\mu\, i\partial_\mu\bar{\chi} = m\psi\,.$$

Before we solve these, note that using one in the other we have

$$\bar{\sigma}^\mu\, i\partial_\mu(\sigma^\nu\, i\partial_\nu\psi) = m^2\psi$$

Since $\partial_\mu\partial_\nu$ is symmetric in $\mu \leftrightarrow \nu$, we can replace $\bar{\sigma}^\mu\sigma^\nu \to \frac{1}{2}\{\bar{\sigma}^\mu, \sigma^\nu\} = \eta^{\mu\nu}$, where the last step follows from the explicit form of $\bar{\sigma}^\mu$ and σ^ν. Hence the right hand side of the equation above is $i^2\eta^{\mu\nu}\partial_\mu\partial_\nu\psi = -\partial^2\psi$ and we have

$$(\partial^2 + m^2)\psi = 0\,.$$

Each component of ψ satisfies the KG equation. Similarly

$$(\partial^2 + m^2)\bar{\chi} = 0\,.$$

6.4.1 Dirac Spinor

It is convenient to combine the two 2-component spinors into a 4-component *Dirac filed*:

$$\Psi = \begin{pmatrix} \psi \\ \bar{\chi} \end{pmatrix}$$

Let

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

This is a 4×4 matrix in 2×2 block notation. Let also

$$\gamma^0\gamma^\mu = \begin{pmatrix} \sigma^\mu & 0 \\ 0 & \bar\sigma^\mu \end{pmatrix} \quad \textit{i.e.} \quad \gamma^\mu = \begin{pmatrix} 0 & \bar\sigma^\mu \\ \sigma^\mu & 0 \end{pmatrix} \quad \text{and} \quad \gamma^i = \begin{pmatrix} 0 & -\sigma^i \\ \sigma^i & 0 \end{pmatrix}$$

Then

$$\mathcal{L} = \Psi^\dagger\gamma^0\gamma^\mu\, i\partial_\mu\Psi - m\Psi^\dagger\gamma^0\Psi$$

or introducing the shorthand $\overline{\Psi} \equiv \Psi^\dagger\gamma^0$,

$$\mathcal{L} = \overline{\Psi}\gamma^\mu\, i\partial_\mu\Psi - m\overline{\Psi} = \overline{\Psi}(\gamma^\mu\, i\partial_\mu - m)\Psi = \overline{\Psi}(i\not\partial - m)\Psi\,.$$

We have introduced the slash notation: for any vector a_μ define $\not a = a_\mu\gamma^\mu$.

The equation of motion is the famous Dirac equation,

$$(i\not\partial - m)\Psi = 0$$

Now

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}$$

so that $(i\not\partial)(i\not\partial) = -\partial_\mu\partial_\nu\gamma^\mu\gamma^\nu = -\partial^2$ and

$$(\partial^2 + m^2)\Psi = 0$$

which is the statement that all components of Ψ satisfy the KG equation, as it should since the components of ψ and $\bar\chi$ do.

Plane-wave expansion: we use $\alpha = 1, \ldots, 4$ for the index of Ψ,

$$\Psi_\alpha(x) = \int (dk) \sum_{s=1}^{2} \left[\beta_{\vec k,s} u_\alpha^{(s)}(\vec k) e^{-ik\cdot x} + \gamma^\dagger_{\vec k,s} v_\alpha^{(s)} e^{ik\cdot x}\right]$$

where $k^2 = m^2$ and the *Dirac spinors* satisfy

$$(\not k - m)u^{(s)}(\vec k) = 0 \quad \text{and} \quad (\not k + m)v^{(s)}(\vec k) = 0\,.$$

To solve these notice that $k^2 = m^2$ gives $(\not k - m)(\not k + m) = 0$ and $(\not k + m)(\not k - m) = 0$. So take $u(\vec k) = (\not k + m)u_0$ for some u_0 such that $(\not k + m)u_0 \neq 0$. Noice that we have anticipated that there are two independent solutions to each of these equations. For example, if $\vec k = 0$, $k^0 = m$ then

$$\not k + m = m(\gamma^0 + 1) = m \begin{pmatrix} \mathbb{1} & \mathbb{1} \\ \mathbb{1} & \mathbb{1} \end{pmatrix}$$

so

$$u^{(1)} = \sqrt{2m} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \; u^{(2)} = \sqrt{2m} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

where we have used a normalization that will give $u(\vec{k})^\dagger u(\vec{k}) = 2E_{\vec{k}}$, as will be needed for simple anti-commutation relations for $\beta_{\vec{k},s}$ and $\gamma_{\vec{k},s}$.

It will be useful to introduce for any 4×4 matrix Γ the conjugate

$$\overline{\Gamma} = \gamma^0 \Gamma^\dagger \gamma^0$$

Then $\overline{\gamma}^\mu = \gamma^0\gamma^{\mu\dagger}\gamma^0 = \gamma^\mu$ (since $\gamma^{0\dagger} = \gamma^0$ and $\gamma^{i\dagger} = -\gamma^i$ and the anti-commutation relations that give $(\gamma^0)^2 = \mathbb{1}$ and $\gamma^0\gamma^i = -\gamma^i\gamma^0$). Then

$$(\not{k} - m)u = 0 \quad \Rightarrow \quad \overline{u}(\not{k} - m) = 0\,,$$

and likewise $\overline{v}(\not{k} + m) = 0$.

It is easy to show that

$$\overline{u}^{(s)}(k)u^{(s')}(k) = 2m\delta^{ss'} = -\overline{v}^{(s)}(k)v^{(s')}(k)$$

and

$$\sum_s u^{(s)}(k)\overline{u}^{(s)}(k) = m + \not{k} \qquad -\sum_s v^{(s)}(k)\overline{v}^{(s)}(k) = m - \not{k} \tag{6.12}$$

Moreover,

$$\overline{u}^{(s)}(k)\gamma^\mu u^{(s)}(k) = 2k^\mu$$

In particular $\overline{u}^{(s)}(k)\gamma^0 u^{(s)}(k) = u^{(s)\dagger}(k)u^{(s)}(k) = 2E$ is not a scalar.

With these normalizations,

$$\{\Psi^\dagger(x), \Psi(y)\}|_{x^0=y^0} = \mathbb{1}\delta^{(3)}(\vec{x} - \vec{y})$$

and

$$H = \sum_s \int (dk) E_{\vec{k}} \left(\beta^\dagger_{\vec{k},s}\beta_{\vec{k},s} + \gamma^\dagger_{\vec{k},s}\gamma_{\vec{k},s}\right)$$

up to an infinite constant, removed by normal-ordering.

6.4.2 Dirac vs Weyl representations

We can always make a redefinition of the Dirac field $\Psi \to S\Psi$ by a unitary matrix S. Then we change $\gamma^\mu \to S^\dagger\gamma^\mu S$. This allows us to choose a different, convenient basis of Dirac gamma matrices. For example we take

$$S = \frac{1}{\sqrt{2}}\begin{pmatrix} \mathbb{1} & \mathbb{1} \\ -\mathbb{1} & \mathbb{1} \end{pmatrix}$$

In this *DIrac representation* we have

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \quad \gamma^0 = \begin{pmatrix} 0 & -\sigma^i \\ \sigma^i & 0 \end{pmatrix}.$$

The basis we had before is called the *Weyl representation*:

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \quad \gamma^0 = \begin{pmatrix} 0 & -\sigma^i \\ \sigma^i & 0 \end{pmatrix}.$$

In any basis,

$$\gamma^{0\dagger} = \gamma^0, \gamma^{i\dagger} = -\gamma^i, \{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}$$

and

$$\mathrm{Tr}\,\not{a} = 0 \tag{6.13}$$

$$\mathrm{Tr}\,\not{a}\not{b} = 4a \cdot b \tag{6.14}$$

$$\mathrm{Tr}\,\not{a}_1 \cdots \not{a}_{2n+1} = 0 \tag{6.15}$$

$$\mathrm{Tr}\,\not{a}\not{b}\not{c}\not{d} = 4(a \cdot b c \cdot d + a \cdot d b \cdot c - a \cdot c b \cdot d) \tag{6.16}$$

6.4.3 Wick's Theorem, T-product, Perturbation theory

Take $x_1^0 > x_2^0$, $\psi = \psi^{(+)} + \psi^{(-)}$, with $\psi^{(+)}$ and $\psi^{(-)\dagger}$ annihilation operators. Then

$$\begin{aligned}\psi(x_1)\psi^\dagger(x_2) &= (\psi_1^{(+)} + \psi_1^{(-)})(\psi_2^{(+)\dagger} + \psi_2^{(-)\dagger}) \\ &= \psi_1^{(+)}\psi_2^{(-)\dagger} + \{\psi_1^{(+)}, \psi_2^{(+)\dagger}\} - \psi_2^{(+)\dagger}\psi_1^{(+)} + \psi_1^{(-)}\psi_2^{(+)\dagger} + \psi_1^{(-)}\psi_2^{(-)\dagger} \\ &= :\psi(x_1)\psi^\dagger(x_2): + c\text{-number}\end{aligned} \tag{6.17}$$

with the understanding that in the normal ordering we pick up a minus sign any time we move an operator through another. So we define the T-ordered product for two anti-commuting fields $A(x)$ and $B(y)$ as

$$T(A(x)B(y)) = \theta(x^0 - y^0)A(x)B(y) - \theta(y^0 - x^0)B(y)A(x)\,.$$

Then the c-number is $\langle 0|T\psi(x_1)\psi^\dagger(x_2)|0\rangle = \overline{\psi(x_1)\psi^\dagger}(x_2)$ and Wick's theorem is just as before with the caveat that we must include minus signs for anti-commutations. For example,

$$\begin{aligned}&\langle 0|T\,\psi(x_1)\psi(x_2)\psi^\dagger(x_3)\psi^\dagger(x_4)|0\rangle \\ &\qquad = \overline{\psi(x_1)\psi^\dagger}(x_4)\overline{\psi(x_2)\psi^\dagger}(x_3) - \overline{\psi(x_1)\psi^\dagger}(x_3)\overline{\psi(x_2)\psi^\dagger}(x_4)\end{aligned} \tag{6.18}$$

The basic quantity we will need to compute amplitudes and for our Feynman rules is the two point function:

$$\langle 0|T\,\Psi_\alpha(x)\overline{\Psi}_\beta(y)|0\rangle = \int \frac{d^4k}{(2\pi)^4} e^{-ik\cdot(x-y)} i \frac{(\not{k} + m)_{\alpha\beta}}{k^2 - m^2 + i\epsilon}$$

Note this is not symmetric under $k \to -k$. You can verify this by writing explicitly the plane-wave expansion of the Dirac field (if you try this you will need to use (6.12)).

To understand Feynman rules we work in a specific context. Let ψ be a Dirac spinor of mass m and ϕ a real scalar of mass M, and take

$$\mathcal{L} = \bar{\psi}(i\partial\!\!\!/ - m)\psi + \tfrac{1}{2}(\partial_\mu\phi)^2 - \tfrac{1}{2}M^2\phi^2 - g\phi\bar{\psi}\psi$$

The last term is called a *Yukawa interaction* and the coefficient g a *Yukawa coupling constant*. Then, as before, Green functions are

$$G^{(n,m,l)}(x_1,\ldots,x_n,y_1,\ldots,y_m,z_1,\ldots,z_l) = \langle 0|T\left(\psi(x_1)\cdots\psi(x_n)\bar{\psi}(y_1)\cdots\bar{\psi}(y_m)\phi(z_1)\cdots\phi(z_l)\right)|0\rangle$$

and in perturbation theory this equals

$$\frac{\langle 0|T\left(\psi_{\rm in}(x_1)\cdots\phi_{\rm in}(z_l)e^{-i\int d^4x g\phi_{\rm in}(x)\bar{\psi}_{\rm in}(x)\psi_{\rm in}(x)}\right)|0\rangle}{\langle 0|T\left(e^{-i\int d^4x g\phi_{\rm in}(x)\bar{\psi}_{\rm in}(x)\psi_{\rm in}(x)}\right)|0\rangle}$$

This can be expanded using Wick's theorem as above. But note that $[\phi,\psi]=0=[\phi,\psi^\dagger]$ so there is no sign change when moving ψ or $\psi^\dagger$ through ϕ.

For example, the simplest non-trivial Green function is $G^{(1,1,1)}$, which to lowest order in an expansion in g is

$$\begin{aligned}G^{(1,1,1)}_{\alpha\beta}(x,y,z) &= -ig\int d^4w\langle 0|T\psi_{{\rm in}\alpha}(x)\psi^\dagger_{{\rm in}\beta}(y)\phi_{\rm in}(z)\phi_{\rm in}(w)\bar{\psi}_{{\rm in}\gamma}(w)\psi_{{\rm in}\gamma}(w)|0\rangle\\ &= -ig\int d^4w(-1)^2\overbrace{\psi_{{\rm in}\alpha}(x)\bar{\psi}_{{\rm in}\gamma}(w)}\overbrace{\psi_{{\rm in}\gamma}(w)\bar{\psi}_{{\rm in}\beta}(y)}\overbrace{\phi_{\rm in}(z)\phi_{\rm in}(w)}\end{aligned}$$

and the rest as before. In particular $G(\{x\}) = \int\prod d^4k e^{i\sum k\cdot x}(2\pi)^4\delta^{(4)}(\sum k)\widetilde{G}(\{k\})$:

$$\begin{aligned}G^{(1,1,1)} &= -ig\int d^4w\int\frac{d^4k}{(2\pi)^4}e^{-ik\cdot(x-w)}\frac{i(k\!\!\!/+m)_{\alpha\gamma}}{k^2-m^2+i\epsilon}\\ &\qquad\times\int\frac{d^4p}{(2\pi)^4}e^{-ip\cdot(w-y)}\frac{i(p\!\!\!/+m)_{\gamma\beta}}{p^2-m^2+i\epsilon}\int\frac{d^4q}{(2\pi)^4}e^{iq\cdot(z-w)}\frac{i}{q^2-M^2+i\epsilon}\\ &= -ig\int\frac{d^4k}{(2\pi)^4}\int\frac{d^4p}{(2\pi)^4}\int\frac{d^4q}{(2\pi)^4}e^{-ik\cdot x+ip\cdot y-iq\cdot z}(2\pi)^4\delta^{(4)}(k-p+q)\\ &\qquad\times i^2\frac{[(k\!\!\!/+m)(k\!\!\!/+m)]_{\alpha\beta}}{(k^2-m^2+i\epsilon)(p^2-m^2+i\epsilon)}\frac{i}{q^2-M^2+i\epsilon}\end{aligned}$$

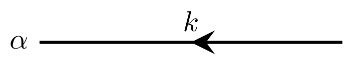
so that

$$\widetilde{G}^{(1,1,1)}(-k,p,-q) = -ig\frac{i}{k\!\!\!/-m}\frac{i}{p\!\!\!/-m}\frac{i}{q^2-M^2}$$

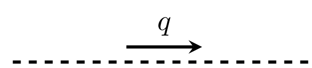
where we have omitted the $i\epsilon$ and used

$$\frac{1}{\not{k} - m} = \frac{\not{k} + m}{k^2 - m^2}\,.$$

We can represent this graphically as follows:

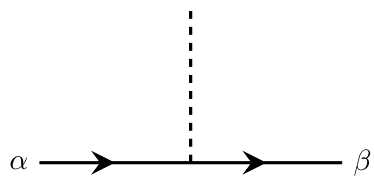


$$= \left(\frac{i}{\not{k} - m + i\epsilon}\right)_{\alpha\beta}$$



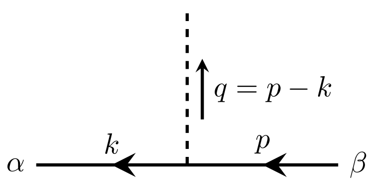
$$= \frac{i}{q^2 - M^2 + i\epsilon}$$

and for the vertex



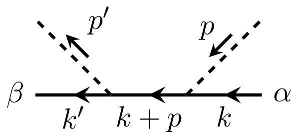
$$= -ig\delta_{\alpha\beta}\,.$$

So the computation above is



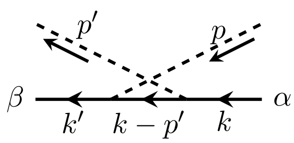
$$= \left(\frac{i}{\not{k} - m + i\epsilon}\right)_{\alpha\gamma} (-ig\delta_{\gamma\delta}) \left(\frac{i}{\not{p} - m + i\epsilon}\right)_{\delta\beta} \frac{i}{q^2 - M^2 + i\epsilon}$$

Here is an example relevant to $\phi\psi \to \phi\psi$ scattering (or rather scattering of the quanta of these fields). To lowest order in an expansion in the coupling constant, the contributions to $\widetilde{G}^{(1,1,2)}$ are



$$= (-ig)^2 \left[\frac{i}{\not{k}' - m}\,\frac{i}{(\not{k} + \not{p}) - m}\,\frac{i}{\not{k} - m}\right]_{\beta\alpha} \frac{i}{p^2 - M^2}\,\frac{i}{p'^2 - M^2}$$

and



$$= (-ig)^2 \left[\frac{i}{\not{k}' - m}\,\frac{i}{(\not{k} - \not{p}') - m}\,\frac{i}{\not{k} - m}\right]_{\beta\alpha} \frac{i}{p^2 - M^2}\,\frac{i}{p'^2 - M^2}$$

6.4.4 LSZ-reduction for spinros, stated

The LSZ reduction formula for spinors is exactly as for scalars, except

(i) Amputate with $i\not\partial - m$ rather than $\partial^2 + m^2$

(ii) An "in" 1-particle state $|\vec{k}, s\rangle_{\rm in}$ gives $u_\alpha^{(s)}(\vec{k})$, where α is contracted with the corresponding index in the Green's function. This comes from

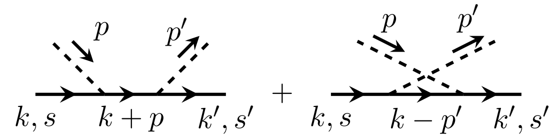
$$\psi_\alpha^{(+)}|\vec{k}, s\rangle_{\rm in} = \int (dk') \sum_{s'} e^{-ik'\cdot x} u_\alpha^{(s')}(\vec{k}') \beta_{\vec{k}', s'} |\vec{k}, s\rangle_{\rm in} .$$

And, similarly, $\bar{u}^{(s)}(\vec{k})$ for ${}_{\rm out}\langle\vec{k}, s|$, $\bar{v}^{(s)}(\vec{k})$ for antiparticle $|\vec{k}, s\rangle_{\rm in}$, and $v^{(s)}(\vec{k})$ for antipartilce ${}_{\rm out}\langle\vec{k}, s|$.

(iii) Possibly $(-1)^p$ for some p, from anti-commuting states.

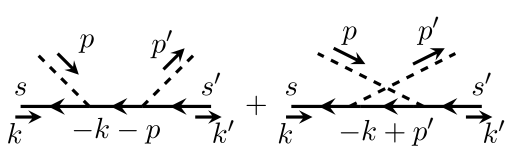
Example: in the sample Yukawa theory above, from the computation of $\widetilde{G}^{(1,1,2)}$, we can obtain the amplitude for the scattering of a spin-0 particle with a spin-$\frac{1}{2}$ particle:

${}_{\rm out}\langle\vec{k}', s'; \vec{p}'|\vec{k}, s; \vec{p}\rangle_{\rm in} =$



$$= -ig^2 \bar{u}^{(s')}(\vec{k}') \left[\frac{1}{(\not k + \not p) - m} + \frac{1}{(\not k - \not p') - m} \right] u^{(s)}(\vec{k})$$

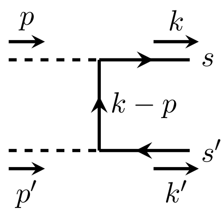
Note the convention here: while time ir ordered later to earlier as we read left to right in ${}_{\rm out}\langle\vec{k}', s'; \vec{p}'|\vec{k}, s; \vec{p}\rangle_{\rm in}$, the Feynman diagram for the amplitude is ordered earlier to later (in to out) as we read left to right. But the expression for the amplitude is ordered, in this case, in the opposite sense: the Dirac spinor for the out state is on the left while the one for the in state is on the right. Generally a line representing a spinor that eneters the diagram from the left and has an arrow pointing right represents an in-particle, while if the arrow is pointing left it represent an in-antiparticle. A line exiting on the right represents an out-particle if the arrow is pointing right, and an out-antiparticle if pointing left. Here is an example of antiparticle scattering off the scalar:



$$= -ig^2 \bar{v}^{(s)}(\vec{k}) \left[\frac{1}{(-\not k - \not p) - m} + \frac{1}{(-\not k + \not p') - m} \right] v^{(s')}(\vec{k}')$$

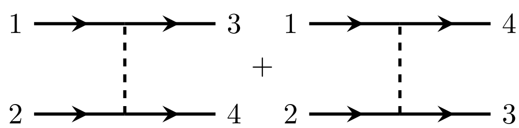
Note that while the particle and antiparticle scattering amplitudes of the scalar appear superficially different, they are the same (up to a sign) once $\not{k}u(\vec{k}) = mu(\vec{k})$ and $\not{k}v(\vec{k}) = -mv(\vec{k})$ are used.

Here is an example which involves both u and v spinors: scalar-scalar scattering into particle -antiparticle pair (of fermions):



$$= -ig^2 \overline{u}^{(s)}(\vec{k}) \left[\frac{1}{(\not{k} - \not{p}) - m} \right] v^{(s')}(\vec{k}')$$

And, finally, here is an example with a sign from anti-commuting external states, two particles in the initial state scattering into two particles:



$$= -ig^2 \left[\frac{\bar{u}_3 u_1 \bar{u}_4 u_2}{(k_1 - k_3)^2 - M^2} - \frac{\bar{u}_4 u_1 \bar{u}_3 u_2}{(k_1 - k_4)^2 - M^2} \right]$$

Here we used the compact notation $u_1 = u^{(s_1)}(\vec{k}_1)$, etc. The relative sign between the two terms is a reflection of Dirac statistics of the external states.

6.5 Generators

Let's study the generators of the Lorentz group in the representation of Weyl spinors:

$$M = 1 - \frac{i}{2}\omega_{\mu\nu}\mathcal{M}^{\mu\nu}$$

corresponding to $\Lambda^{\mu\nu} = \eta^{\mu\nu} + \omega^{\mu\nu}$ with $\omega^{\mu\nu} = -\omega^{\nu\mu}$ infinitesimal. So $\mathcal{M}^{\mu\nu}$ are six 2×2 matrices that we want to characterize. First we note that $\det(M) = 1$ implies $\mathrm{Tr}(\mathcal{M}^{\mu\nu}) = 0$. Next, determine $\mathcal{M}^{\mu\nu}$, using the known transformation properties of vectors. On one hand

$$\begin{aligned} M^\dagger P M &= ((\Lambda^{-1})_\mu{}^\nu p_\nu)\sigma^\mu \\ &= p^\nu \sigma^\mu (\eta_{\mu\nu} - \omega_{\mu\nu}) \end{aligned}$$

and on the other

$$\begin{aligned} M^\dagger P M &= (1 + \tfrac{i}{2}\omega_{\mu\nu}\mathcal{M}^{\dagger\mu\nu}) p_\lambda \sigma^\lambda (1 - \tfrac{i}{2}\omega_{\sigma\rho}\mathcal{M}^{\sigma\rho}) \\ &= p_\mu \sigma^\mu + \tfrac{i}{2}\omega_{\mu\nu} p_\lambda (\mathcal{M}^{\dagger\mu\nu}\sigma^\lambda - \sigma^\lambda \mathcal{M}^{\mu\nu}) \, . \end{aligned}$$

Equating,

$$\tfrac{i}{2}\omega_{\mu\nu}p_\lambda(\mathcal{M}^{\dagger\mu\nu}\sigma^\lambda - \sigma^\lambda\mathcal{M}^{\mu\nu}) = -p^\nu\sigma^\mu\omega_{\mu\nu} = -\omega_{\mu\nu}p_\lambda\eta^{\lambda\nu}\sigma^\mu$$

and since this must hold for arbitrary $\omega_{\mu\nu}$ and p_λ we have

$$\mathcal{M}^{\dagger\mu\nu}\sigma^\lambda - \sigma^\lambda\mathcal{M}^{\mu\nu} = -i(\eta^{\lambda\mu}\sigma^\nu - \eta^{\lambda\nu}\sigma^\mu), .$$

The solution to this is straightforward:

$$\mathcal{M}^{0i} = -\mathcal{M}^{i0} = \tfrac{i}{2}\sigma^i, \quad \mathcal{M}^{ij} = -\mathcal{M}^{ji} = \tfrac{1}{2}\epsilon^{ijk}\sigma^k,$$

(Aside: to determine this, expand $\mathcal{M}^{\mu\nu}$ in the basis of σ^μ. Since $\mathrm{Tr}\,\mathcal{M}^{\mu\nu} = 0$ we have $\mathcal{M}^{\mu\nu} = (a_j^{\mu\nu} + ib_j^{\mu\nu})\sigma^j$ and $\mathcal{M}^{\dagger\mu\nu} = (a_j^{\mu\nu} - ib_j^{\mu\nu})\sigma^j$. This gives

$$a_j^{\mu\nu}[\sigma^j, \sigma^\lambda] - ib_j^{\mu\nu}\{\sigma^j, \sigma^\lambda\} = -i(\eta^{\lambda\mu}\sigma^\nu - \eta^{\lambda\nu}\sigma^\mu)\,.$$

Setting $\lambda = 0$ gives an equation for the b_j's:

$$2b_j^{\mu\nu}\sigma^j = -(\delta^{\nu 0}\sigma^\mu - \delta^{\mu 0}\sigma^\nu)$$

and taking $\mu = 0$ and $\nu = k$ we have $2b_j^{0k}\sigma^j = \sigma^k$ from which $b_j^{0k} = -b_j^{k0} = \frac{1}{2}\delta_j^k$ follows. Similarly we obtain $b_j^{ik} = 0$. For the $a's$ set $\lambda = i$. Then

$$a_j^{\mu\nu}2i\epsilon^{jik}\sigma^k - ib_j^{\mu\nu}2\delta^{ij} = -i(\eta^{i\mu}\sigma^\nu - \eta^{i\nu}\sigma^\mu)$$

Then setting $\mu = 0$ gives $a_j^{0\nu} = 0$ and setting $\mu = l$ and $\nu = m$ gives $a_j^{lm} = \frac{1}{2}\epsilon^{lmj}$.)

Let

$$J^k = \tfrac{1}{2}\epsilon^{kij}\mathcal{M}^{ij} = \tfrac{1}{2}\sigma^k, \quad \text{and} \quad K^i = \mathcal{M}^{0i} = \tfrac{i}{2}\sigma^i\,.$$

Note that $J^{i\dagger} = J^i$ while $K^{i\dagger} = -K^i$. These satisfy,

$$[J^i, J^j] = i\epsilon^{ijk}J^k, \quad [J^i, K^j] = i\epsilon^{ijk}K^k, \quad [K^i, K^j] = -i\epsilon^{ijk}J^k\,.$$

Defining

$$J_\pm^i \equiv \tfrac{1}{2}(J^i \pm iK^i)$$

we have

$$[J_+^i, J_+^j] = i\epsilon^{ijk}J_+^k, \quad [J_-^i, J_-^j] = i\epsilon^{ijk}J_-^k, \quad [J_+^i, J_-^j] = 0\,.$$

You recognize these are two mutually commuting copies of (the algebra of) $SO(3) \sim SU(2)$. You know this from the rotation group in QM: for spinors, $\vec{S} = \frac{1}{2}\vec{\sigma}$, and $[S^i, S^j] = i\epsilon^{ijk}S^k$, while for vectors $(L^i)_{jk} = -i\epsilon^{ijk}$, with $[L^i, L^j] = i\epsilon^{ijk}L^k$. The irreducible representations of $SU(2)$ are $\ell = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$ of dimension $2\ell + 1 = 1, 2, 3, 4, \ldots$. You recognize $\ell = 0$ as a scalar, $\ell = \frac{1}{2}$ a spinor, $\ell = 1$ a vector, etc.

Comments:

(i) For any representation $D(\Lambda)$ we have infinitesimal generators. If $\Lambda^{\mu\nu} = \eta^{\mu\nu} + \omega^{\mu\nu}$ then $D(\Lambda)_{\alpha\beta} = \delta_{\alpha\beta} - \frac{i}{2}\omega_{\mu\nu}\mathcal{J}^{\mu\nu}$. The six matrices $\mathcal{J}^{\mu\nu}$ satisfy the same commutation relations as $\mathcal{M}^{\mu\nu}$, they are fixed by the multiplication "table" of the Lorentz group, which itself follows form requiring $D(\Lambda)$ be a representation, Eq. (6.2).

(ii) The *defining* representation is the 4-dim representation acting on vectors, p^μ, like Λ itself. That is, $D(\Lambda)^\mu{}_\nu p^\nu = \Lambda^\mu{}_\nu p^\nu \Rightarrow -\frac{i}{2}\omega_{\lambda\sigma}(\mathcal{J}^{\lambda\sigma})_{\mu\nu}p^\nu = \omega_{\mu\nu}p^\nu$, from which we read off, for the 4-dimensional (defining) representation of the Lorentz group: $(\mathcal{J}^{\lambda\sigma})_{\mu\nu} = -i(\delta^\lambda_\mu\delta^\sigma_\nu - \delta^\sigma_\mu\delta^\lambda_\nu)$. We can compute easily the same commutation relations satisfied by the $\mathcal{J}^{\lambda\sigma}$ and of course they are the same as those satisfied by $\mathcal{M}^{\lambda\sigma}$.

(iii) The commutation relations for J^i and K^i derived above may seem ambiguous since they were found from comparing with Pauli matrices but both J^i and K^i are given in terms of Pauli matrices. For example, $[K^i, K^j] = -i\epsilon^{ijk}\frac{1}{2}\sigma^k$ was written as $-i\epsilon^{ijk}J^k$ rather than $-\epsilon^{ijk}K^k$. There is a simple argument why $[K, K] = -K$ is excluded. Under parity $\mathcal{M}^{ij} \to (-1)^2\mathcal{M}^{ij}$, $\mathcal{M}^{0i} \to (-1)\mathcal{M}^{0i}$. Hence $J^i \to J^i$ and $K^i \to -K^i$. This gives, $[J, J] \sim J$ but not K, $[J, K] \sim K$ but not J and $[K, K] \sim J$ but not K.

(iv) We would not have faced this ambiguity (nor would we have had to use the parity argument to sort it out) had we studied the commutation relations for arbitrary representations.

6.5.1 All the representations of the Lorentz Group

Since the same commutation relations must hold for any representation we use that to construct them. We build on our knowledge of the representations of $SU(2)$.

For J^i_+ the irreducible representations are classified by $s_+ = 0, \frac{1}{2}, 1, \ldots$. The $(2s_+{+}1)\times(2s_+{+}1)$ matrices are labeled $J^i_{s_+}$. For J^i_- the irreducible representations are classified by $s_- = 0, \frac{1}{2}, 1, \ldots$. The $(2s_-{+}1)\times(2s_-{+}1)$ matrices are labeled $J^i_{s_-}$. Moreover, since the representation matrices of the generators of the Lorentz group, J^i_+ and J^i_-, satisfy $[J^i_+, J^i_-] = 0$, they are tensor products, $J^i_+ = J^i_{s_+} \otimes \mathbb{1}_{s_-}$, where $\mathbb{1}_{s_-}$ is the $(2s_- + 1) \times (2s_- + 1)$ identity, matrix, and, similarly, $J^i_- = \mathbb{1}_{s_+} \otimes J^i_{s_-}$. The irreducible representation are labeled by (s_+, s_-) and have generators $J^i = J^i_+ + J^i_- = J^i_{s_+} \otimes \mathbb{1}_{s_-} + \mathbb{1}_{s_+} \otimes J^i_{s_-}$ and similarly for K^i. They have dimension $(2s_+ + 1)(2s_- + 1)$.

For example, $(0, 0)$ is a representation of dimension 1, a scalar.

$(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$ have dimension 2. They are spinors. They are two different spinor representations. Consider $(\frac{1}{2}, 0)$:

$$J^i_{s_+} = \tfrac{1}{2}\sigma^i,\ J^i_{s_-} = 0,\ \ \Rightarrow \vec{J} = \vec{J}_+ + \vec{J}_- = \tfrac{1}{2}\vec{\sigma} \otimes \mathbb{1},\ \vec{K} = \tfrac{1}{i}(\vec{J}_+ - \vec{J}_-) = -\tfrac{i}{2}\vec{\sigma} \otimes \mathbb{1}$$

Similarly, for $(0, \frac{1}{2})$

$$J^i_{s_+} = 0,\ J^i_{s_-} = \tfrac{1}{2}\sigma^i,\ \ \Rightarrow \vec{J} = \vec{J}_+ + \vec{J}_- = \mathbb{1}\otimes\tfrac{1}{2}\vec{\sigma},\ \vec{K} = \tfrac{1}{i}(\vec{J}_+ - \vec{J}_-) = \mathbb{1}\otimes\tfrac{i}{2}\vec{\sigma}$$

But this is where we started from. From $\psi' = M\psi$ we obtained $\vec{J} = \frac{1}{2}\vec{\sigma}$ and $\vec{K} = \frac{i}{2}\vec{\sigma}$ so this is the $(0, \frac{1}{2})$ representation. So what is $(\frac{1}{2}, 0)$? We already know that $\overline{M}$ is a representation. Look more closely:

$$M = e^{-i\vec{\theta}\cdot(\frac{1}{2}\vec{\sigma}) - i\vec{\gamma}\cdot(\frac{i}{2}\vec{\sigma})} = e^{-i\vec{\alpha}\cdot(\frac{1}{2}\vec{\sigma})} \quad \text{with } \vec{\alpha} = \vec{\theta} + i\vec{\gamma},\ (\alpha^i \in \mathbb{C}).$$

The representation we are looking for should have group elements

$$e^{-i\vec{\theta}\cdot(\frac{1}{2}\vec{\sigma}) - i\vec{\gamma}\cdot(-\frac{i}{2}\vec{\sigma})} = e^{-i\vec{\alpha}^*\cdot(\frac{1}{2}\vec{\sigma})}$$

while

$$M^* = e^{-i\vec{\theta}\cdot(-\frac{1}{2}\vec{\sigma}^*) + i\vec{\gamma}\cdot(-\frac{i}{2}\vec{\sigma}^*)} = e^{-i\vec{\alpha}^*\cdot(-\frac{1}{2}\vec{\sigma}^*)}\,.$$

This is close. In fact, it is what we want up to a similarity transformation: using $\sigma^2\vec{\sigma}^*\sigma^2 = -\vec{\sigma}$ we have

$$\overline{M} = \sigma^2 M^* \sigma^2 = e^{-i\vec{\alpha}^*\cdot(\frac{1}{2}\vec{\sigma})}\,.$$

That is, M^* is in the equivalence class of $(\frac{1}{2}, 0)$.

Generalize: note that since $M = \mathbb{1} - i\omega_{\mu\nu}\mathcal{M}^{\mu\nu}$ we have $M^* = \mathbb{1} - i\omega_{\mu\nu}(-\mathcal{M}^{\mu\nu *})$. The matrices $-\mathcal{M}^{\mu\nu *}$ satisfy the same commutation relations as the $\mathcal{M}^{\mu\nu}$. More genrally, if $[T^a, T^b] = if^{abc}T^c$ then $[-T^{a*}, -T^{b*}] = if^{abc*}(-T^{c*})$, so $-T^{a^*}$ satisfy the same commutation relations as T^a if f^{abc} are real. In our case ($J^i_\pm$) the f^{abc} are ϵ^{ijk}. Since for $SU(2)$ the only irreducible representation of dimension $2s+1$ is generated by $\vec{J}_s$ it must be that $S(-\vec{J}^*_s)S^{-1} = \vec{J}_s$ for some invertible matrix S. Now

$$D(\Lambda) = e^{-i\vec{\alpha}\cdot\vec{J}_+ - i\vec{\alpha}^*\cdot\vec{J}_-}$$

so that

$$SD^*(\Lambda)S^{-1} = e^{-i\vec{\alpha}^*\cdot\vec{J}_+ - i\vec{\alpha}\cdot\vec{J}_-}$$

The role of J_+ and J_- has been exchanged. To be more precise, the matrix S that acts on the tensor product exchanges the $+$ and $-$ sectors. We therefore have,

$$(s_+, s_-)^* \sim (s_-, s_+)$$

Note that for $s_+ = s_-$ the complex conjugate representation is similar to itself. This is a *real representation* and the vectors on which it acts can be taken to have real components. For example, let's investigate the $(\frac{1}{2}, \frac{1}{2})$ representation. It is 2×2 dimensional. It smells like a 4-vector. Let's show it is. It is an object with indices $\dot{\alpha}\alpha$ as in $V_{\dot{\alpha}\alpha}$, with

$$V'_{\dot{\alpha}\alpha} = \overline{M}_{\dot{\alpha}\dot{\beta}} M_{\alpha\beta} V_{\dot{\beta}\beta}$$

This transforms like $\bar{\chi}_{\dot{\alpha}}\psi_\alpha$. We have already seen that this 4-component object can be arranged into a 4-vector; more specifically $\psi^\dagger\sigma^\mu\psi$ is a 4-vector and the relation between $\psi^\dagger = \bar{\chi}^\dagger\sigma^2$, so consider $V^\mu = V_{\dot{\alpha}\alpha}(\sigma^2\sigma^\mu)_{\dot{\alpha}\alpha}$. Then

$$\begin{aligned}
V'^\mu &= \overline{M}_{\dot{\alpha}\dot{\beta}} M_{\alpha\beta} V_{\dot{\beta}\beta}(\sigma^2\sigma^\mu)_{\dot{\alpha}\alpha} \\
&= (\overline{M}^T\sigma^2\sigma^\mu M)_{\dot{\beta}\beta} V_{\dot{\beta}\beta} \\
&= (\sigma^2 M^\dagger\sigma^\mu M)_{\dot{\beta}\beta} V_{\dot{\beta}\beta} \\
&= \Lambda^\mu{}_\nu(\sigma^2\sigma^\nu)_{\dot{\beta}\beta} V_{\dot{\beta}\beta} \\
&= \Lambda^\mu{}_\nu V^\nu
\end{aligned}$$

More generally, $X_{\alpha_1\cdots\alpha_{s_+}\dot{\alpha}_1\cdots\alpha_{s_-}}$ with all α indices symmetrized and all $\dot{\alpha}$ indices symmetrized is in the (s_+, s_-) representation.