Quantum Field Theory: An Introduction

Ryan Reece[1]

[1]ryan.reece@cern.ch, reece.scipp.ucsc.edu,
Santa Cruz Institute for Particle Physics, University of California,
1156 High St., Santa Cruz, CA 95064, USA

December 23, 2007

Abstract

This document is a set of notes I took on QFT as a graduate student at the University of Pennsylvania, mainly inspired in lectures by Burt Ovrut, but also working through Peskin and Schroeder (1995), as well as David Tong's lecture notes available online. They take a slow pedagogical approach to introducing classical field theory, Noether's theorem, the principles of quantum mechanics, scattering theory, and culminating in the derivation of Feynman diagrams.

Contents

1 Preliminaries

1.1 Overview of Special Relativity

1.1.1 Lorentz Boosts

Searches in the later part 19th century for the coordinate transformation that left the form of Maxwell's equations and the wave equation invariant lead to the discovery of the Lorentz Transformations. The "boost" transformation from one (unprimed) inertial frame to another (primed) inertial frame moving with dimensionless velocity $\vec{\beta} = \vec{v}/c$, respect to the former frame, is given by

$$\begin{pmatrix} c\,t' \\ x' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} c\,t \\ x \end{pmatrix}$$

Because a boost along one of the spacial dimensions leaves the other two unchanged, we can suppress the those two spacial dimensions and let $\beta = |\vec{\beta}\,|$. γ is the **Lorentz Factor**, defined by

$$\gamma \equiv \frac{1}{\sqrt{1-\beta^2}} \tag{1}$$

γ ranges from 1 to ∞ monotonicly in the nonrelativistic ($\beta \to 0$) and relativistic ($\beta \to 1$) limits, respectively. It is useful to remember that $\gamma \geq 1$. Note that being the magnitude of a vector, β has a lower limit at 0. β also has an upper limit at 1 because γ diverges as β approaches 1 and becomes unphysically imaginary for values of $\beta > 1$. This immediately reveals that $\beta = 1$, or $v = c$, is Nature's natural speed limit.

The inverse transformation is given by

$$\begin{pmatrix} c\,t \\ x \end{pmatrix} = \begin{pmatrix} \gamma & \gamma\beta \\ \gamma\beta & \gamma \end{pmatrix} \begin{pmatrix} c\,t' \\ x' \end{pmatrix}$$

1.1.2 Length Contraction and Time Dilation

The differences between two points in spacetime follow from the transformations:

$$\begin{aligned} c\,\Delta t' &= \gamma\,c\,\Delta t - \gamma\beta\,\Delta x \\ \Delta x' &= -\gamma\beta\,c\,\Delta t + \gamma\,\Delta x \\ c\,\Delta t &= \gamma\,c\,\Delta t' + \gamma\beta\,\Delta x' \\ \Delta x &= \gamma\beta\,c\,\Delta t' + \gamma\,\Delta x' \end{aligned}$$

Consider a clock sitting at rest in the unprimed frame ($\Delta x = 0$). The first of the four above equations and the fact that $\gamma \geq 1$, imply that the *time interval is dilated* in the primed frame.

$$\Delta t' = \gamma\,\Delta t$$

Now consider a rod of length Δx in the unprimed frame. A measurement of the length in the primmed frame corresponds to determining the coordinates of the endpoints simultaneously in the unprimed frame ($\Delta t' = 0$). Then the fourth equation implies that *length is contracted* in the primed frame.

$$\Delta x' = \frac{\Delta x}{\gamma}$$

We call time intervals and lengths "proper" if they are measured in the frame where the subject is at rest (in this case, the unprimed frame). In summary, proper times and lengths are the shortest and longest possible, respectively.

1.1.3 Four-vectors

Knowing that lengths and times transform from one reference frame to another, one can wonder if there is anything that is invariant. Consider the following, using the last two of the four equations for the differences between two spacetime points.

$$\begin{aligned}
(c\,\Delta t)^2 - (\Delta x)^2 &= \left(\gamma\,c\,\Delta t' + \gamma\,\beta\,\Delta x'\right)^2 - \left(\gamma\beta\,c\,\Delta t' + \gamma\,\Delta x'\right)^2 \\
&= \gamma^2\left[(c\,\Delta t')^2 + \cancel{2\beta\,c\,\Delta t'\,\Delta x'} + \beta^2\,(\Delta x')^2\right. \\
&\qquad \left. -\beta^2\,(c\,\Delta t')^2 - \cancel{2\beta\,c\,\Delta t'\,\Delta x'} - (\Delta x')^2\right] \\
&= \gamma^2\,\underbrace{\left(1-\beta^2\right)}_{\gamma^{-2}}\,\left[(c\,\Delta t')^2 - (\Delta x')^2\right] \\
&= (c\,\Delta t')^2 - (\Delta x')^2 \equiv (\Delta\tau)^2
\end{aligned}$$

Which shows that $\Delta\tau$ has the same value in any frames related by Lorentz Transformations. $\Delta\tau$ is called the "**invariant length**." Note that it is equal to the proper time interval.

This motivates us to think of $(t, \vec{x}\,)$ as a **four-vector** that transforms according to the Lorentz transformations, in a "spacetime vector space," and there should be some kind of "inner product," or contraction, of these vectors that leaves $\Delta\tau$ a scalar. This can be done by defining the Minkowski metric tensor as follows.

$$\eta_{\mu\nu} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}_{\mu\nu} \tag{2}$$

Four-vectors are indexed by a Greek index, $x^\mu = (c\,t, \vec{x}\,)^\mu$, μ ranging from 0 to 3 ($x^0 = c\,t$, $x^1 = x$, $x^2 = y$, $x^3 = z$). The contraction of a spacetime four-vector with itself, its square, is give by

$$x^{\mu}\, x_{\mu} \equiv x^{\mu}\, \eta_{\mu\nu}\, x^{\nu} = (c\, t)^2 - \vec{x} \cdot \vec{x} = (\Delta\tau)^2 \tag{3}$$

giving the square of the invariant length between x^{μ} and the origin. In equation (3), we have defined that the lowering of a four-vector index is done by multiplication by the metric tensor. Explicit matrix multiplication will show that $\eta^{\mu\nu}$ is the inverse Minkowski metric and has the same components as $\eta_{\mu\nu}$.

$$\eta_{\mu\lambda}\, \eta^{\lambda\nu} = \delta_{\mu}^{\nu} \tag{4}$$

Raising the indices of the metric confirms that the components of the metric and inverse metric are equal.

$$\eta^{\mu\nu} = \eta^{\mu\lambda}\, \eta^{\nu\sigma}\, \eta_{\lambda\sigma} = \eta^{\mu\lambda}\, \eta_{\lambda\sigma}\, (\eta^{\mathrm{T}})^{\sigma\nu}$$

Anything that transforms according to the Lorentz Transformations, like $(c\, t, \vec{x}\,)$, is a four-vector. Another example of a four-vector is **four-velocity**, defined by

$$u^{\mu} \equiv \gamma\, (c, \vec{v}\,)^{\mu} \tag{5}$$

One can show that the square of u^{μ} is invariant as required.

$$\begin{aligned} u^{\mu}\, u_{\mu} &= \gamma^2\, (c^2 - v^2) \\ &= \frac{1}{1-\beta^2}\, (1-\beta^2)\, c^2 \\ &= c^2 \end{aligned}$$

which is obviously invariant. Any equation where all of the factors are scalars (with no indices or contracting indices), or are four-vectors/tensors, with matching indices on the other side of the equal sign, is called “manifestly invariant.”

1.1.4 Momentum and Energy

The Classically conserved definitions of momentum and energy, being dependent on the coordinate frame, will not be conserved in other frames. We are motivated to consider the effect of defining momentum with the four-velocity instead of the classical velocity. The mass of a particle, m, being an intrinsic property of the particle, must be a Lorentz scalar. Therefore, the following definition of the **four-momentum** is manifestly a four-vector.

$$p^{\mu} \equiv m\, u^{\mu} = \gamma\, m\, (c, \vec{v}\,)^{\mu} \tag{6}$$

The square of which is

$$p^{\mu}\, p_{\mu} = m^2\, u^{\mu}\, u_{\mu} = m^2\, c^2 \tag{7}$$

Now let's give some interpretation to the components of the four-momentum. To consider the nonrelativistic limit, let us expand γ in the $\beta \to 0$ limit.

$$\gamma \simeq 1 + \frac{1}{2}\,\beta^2 + \frac{3}{8}\,\beta^4 + \cdots$$

Then the leading order term of the space-like components of the four-momentum is just the Classical momentum.

$$\vec{p} = m\,\vec{v} + \cdots$$

We therefore, interpret the space-like components of the four-momentum as the relativistic momentum.

$$\vec{p} = \gamma\, m\,\vec{v} \tag{8}$$

The expansion of the time-like term gives

$$m\,c^2 + \frac{1}{2}\,m\,v^2 + \cdots$$

We can now recognize the second term as the Classical kinetic energy. The first term is evidently the "mass energy," energy present even when $v = 0$. Higher order terms give relativistic corrections.

$$E = \gamma\, m\, c^2 \tag{9}$$

We can therefore write the four-momentum in terms of the relativistic energy, E, and relativistic momentum, p.

$$\boxed{p^\mu = (E, \vec{p})^\mu} \tag{10}$$

The four-momentum is the combination of momentum and energy necessary to transform according to Lorentz Transformations. Both E and $\vec{p}$ are *conserved* quantities in any given frame, but they are *not invariant*; they *transform* when going to another frame. Scalar quantities, like mass, are *invariant* but are *not necessarily conserved.* Mass can be exchanged for kinetic energy and *vice versa.* Charge is an example of a scalar quantity that is also conserved.

Looking at the square of the four-momentum with this energy-momentum interpretation of its components gives the very important relationship between energy, momentum, and mass.

$$\boxed{p^\mu\, p_\mu = E^2 - |\vec{p}|^2\, c^2 = m^2\, c^4} \tag{11}$$

Taking the ratio of equations (8) and (9) gives the following interesting relation.

$$\frac{\vec{p}}{E} = \frac{\vec{v}}{c^2} \tag{12}$$

which leads to

$$\boxed{\frac{\vec{p}\,c}{E} = \vec{\beta}} \tag{13}$$

Note from equation (11), in the case that $m = 0$, we have that $E = |\vec{p}|\,c$, which implies that $\beta = 1$. Therefore, *massless particles must travel at the speed of light.* In which case (13) agrees that the following is the energy-momentum relation for massless particles.

$$E = |\vec{p}|\,c \tag{14}$$

Also note the relationship to Einstein's equation for the energy of a photon, $E = \hbar\omega \;\Rightarrow\; |\vec{p}| = \hbar\,\frac{\omega}{c} = \hbar\,k$, consistent with de Broglie's relation.

1.2 Units

1.2.1 Natural Units

Factors of c were explicit in the above review of special relativity. From now on, we will use a form of **natural units**, where certain natural constants are set to one by using units derived from the God-given scales in Nature.

$$\hbar = c = \varepsilon_0 = 1 \tag{15}$$

From $\hbar = 6.58 \times 10^{-25}$ GeV $\cdot$ s $= 1$, it follows that if we choose to measure energy in units GeV, then time can be measured in units GeV^{-1}.

$$1\ \text{GeV}^{-1} = 6.58 \times 10^{-25}\ \text{s} \tag{16}$$

From $c = 3 \times 10^{8}$ m/s $= 1$, it follows that

$$1 = \left(3 \times 10^{8}\text{m}\right)\left(6.58 \times 10^{-25}\ \text{GeV}\right) = 1.97 \times 10^{-16}\ \text{m} \cdot \text{GeV} \tag{17}$$

$$\Rightarrow \qquad 1\ \text{GeV}^{-1} = 1.97 \times 10^{-16}\ \text{m} \tag{18}$$

Summarizing the dimensionality:

$$\text{time} = \text{length} = \frac{1}{\text{energy}} \tag{19}$$

1.2.2 Barns

When calculating cross sections, the conventional unit of area in particle physics is a **barn**.

$$1\ \text{barn} \equiv (10\ \text{fm})^2 = 10^{-24}\ \text{cm}^2 \tag{20}$$

$$1 \text{ mb} \equiv 10^{-3} \text{ barns} = 10^{-27} \text{ cm}^2 \tag{21}$$

From (18), it can be shown that

$$\boxed{1 \text{ GeV}^{-2} = 0.389 \text{ mb}} \tag{22}$$

1.2.3 Electromagnetism

Finally, from $\varepsilon_0 = 1$ and $c = 1$

$$c = \frac{1}{\sqrt{\mu_0 \, \varepsilon_0}} \qquad \Rightarrow \qquad \mu_0 = 1 \tag{23}$$

giving Maxwell's equations the following form.
Field Tensor:

$$F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu \tag{24}$$

Homogeneous:

$$\partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu} = 0 \tag{25}$$

Inhomogeneous:

$$\partial_\nu F^{\mu\nu} = J^\mu \tag{26}$$

1.3 Relativistic Kinematics

1.3.1 Lorentz Invariant Phase Space

1.3.2 Mandelstam Variables

2 Variation of Fields

2.1 The Field Worldview

We will see that the dynamical variable in quantum field theory is the field itself. A **field** is a mathematical concept that has a number or a construction of numbers (a complex number, vector, spinor, tensor ...) defined at every point in spacetime. We assume that the fields are smoothly varying, such that derivatives are well defined. To simplify notation, in these first few sections, we will denote a general field by $\phi_a(x)$, where a indexes all of the components of ϕ, being components of vectors, spinors, etc. or some direct product of them.

2.2 Variation

By studying the variation of fields we will discover the effect the Principle of Least Action has on dynamics, giving the Euler-Lagrange equation. Then we will discover the effects the symmetries of spacetime have on dynamics, giving Noether's Theorem. But first, we need to derive some properties of variation of functions in general.

We define the following.

$$\begin{aligned} \delta x^\mu &\equiv x'^\mu - x^\mu \\ \delta_o\phi(x) &\equiv \phi'(x) - \phi(x) \\ \delta\phi(x) &\equiv \phi'(x') - \phi(x) \end{aligned}$$

δx^μ is the difference in the values of coordinates referring to the same spacetime point but in different coordinate frames. $\delta_o\phi(x)$ is the value of the field $\phi(x)$ subtracted from the value of a field with a slightly different functional form, $\phi'(x)$, evaluated at the same spacetime point, in the same coordinates. $\delta\phi(x)$ is the difference in the functions $\phi'(x')$ and $\phi(x)$ evaluated in in different coordinate systems.

Using the above definitions we can derive the following.

$$\begin{aligned} \delta\phi(x) &= \phi'(x + \delta x) - \phi(x) \\ &= \phi'(x) + \delta x^\mu\, \partial_\mu\phi'(x) + \ldots - \phi(x) \\ &\simeq \delta_o\phi(x) + \delta x^\mu\, \partial_\mu\phi'(x) \\ &= \delta_o\phi(x) + \delta x^\mu\, \partial_\mu\big(\phi(x) + \delta_o\phi(x)\big) \\ &= \delta_o\phi(x) + \delta x^\mu\, \partial_\mu\phi(x) + \ldots \end{aligned}$$

We have used that the variations are small and that field is smooth by Taylor expanding

$\phi'(x)$. Therefore, to leading order in the variation, we have

$$\boxed{\delta\phi(x) = \delta_o\phi(x) + \delta x^\mu\, \partial_\mu \phi(x)} \tag{27}$$

Similarly for a function of several variables, we have

$$\begin{aligned} \delta_o f(x,y) &\equiv f'(x,y) - f(x,y) \\ \delta f(x,y) &\equiv f'(x',y') - f(x,y) \end{aligned}$$

$$\begin{aligned} \delta f(x,y) &= f'(x+\delta x, y+\delta y) - f(x,y) \\ &= f'(x,y) + \delta x^\mu \frac{\partial}{\partial x^\mu} f'(x,y) + \delta y^\mu \frac{\partial}{\partial y^\mu} f'(x,y) + \ldots - f(x,y) \\ &\simeq \delta_o f(x,y) + \delta x^\mu \frac{\partial}{\partial x^\mu} f'(x,y) + \delta y^\mu \frac{\partial}{\partial y^\mu} f'(x,y) \\ &= \delta_o f(x,y) + \delta x^\mu \frac{\partial}{\partial x^\mu}\big(f(x,y) + \delta_o f(x,y)\big) + \delta y^\mu \frac{\partial}{\partial y^\mu}\big(f(x,y) + \delta_o f(x,y)\big) \\ &= \delta_o f(x,y) + \delta x^\mu \frac{\partial}{\partial x^\mu} f(x,y) + \delta y^\mu \frac{\partial}{\partial y^\mu} f(x,y) + \ldots \end{aligned}$$

$$\therefore \qquad \boxed{\delta f(x,y) = \delta_o f(x,y) + \delta x^\mu \frac{\partial}{\partial x^\mu} f(x,y) + \delta y^\mu \frac{\partial}{\partial y^\mu} f(x,y)} \tag{28}$$

One can begin to see that variation, δ, follows similar operational rules as that of the differential operator, d. Indeed, we can derive a product rule for variation as follows.

$$\text{Let} \qquad f(x,y) \equiv g(x)\, h(y)$$

$$\Rightarrow \qquad \delta f(x,y) = \delta_o\big(g(x)\, h(y)\big) + h(y)\, \delta x^\mu \frac{\partial}{\partial x^\mu} g(x) + g(x)\, \delta y^\mu \frac{\partial}{\partial y^\mu} h(y)$$

$$\begin{aligned} \delta_o\big(g(x)\, h(y)\big) &= g'(x)\, h'(y) - g(x)\, h(y) \\ &= \big(g(x) + \delta_o g(x)\big)\big(h(y) + \delta_o h(y)\big) - g(x)\, h(y) \\ &= \cancel{g(x)\, h(y)} + \big(\delta_o g(x)\big) h(y) + g(x)\, \delta_o h(y) + \cancelto{\mathcal{O}[\delta^2]}{\big(\delta_o g(x)\big)\big(\delta_o h(y)\big)} - \cancel{g(x)\, h(y)} \\ &= \big(\delta_o g(x)\big) h(y) + g(x)\, \delta_o h(y) \end{aligned}$$

$$\Rightarrow \qquad \delta f(x,y) = h(y) \left(\delta_o g(x) + \delta x^\mu \frac{\partial}{\partial x^\mu} g(x)\right) + g(x) \left(\delta_o h(y) + \delta y^\mu \frac{\partial}{\partial y^\mu} h(y)\right)$$

$$\therefore \qquad \boxed{\delta\big(g(x)\, h(y)\big) = h(y)\, \delta g(x) + g(x)\, \delta h(y)} \tag{29}$$

One can show that δ_o commutes with partial derivatives as follows.

$$\begin{aligned}
\partial_\mu \delta_o \phi(x) &= \partial_\mu \big(\phi'(x) - \phi(x)\big) \\
&= \partial_\mu \phi'(x) - \partial_\mu \phi(x) \\
&= \delta_o \partial_\mu \phi(x)
\end{aligned}$$

$$\therefore \qquad \boxed{\delta_o \partial_\mu \phi(x) = \partial_\mu \delta_o \phi(x)} \tag{30}$$

2.3 The Principle of Least Action

The Classical Mechanics of a field can be described by introducing a **Lagrangian density** (often just called a Lagrangian), $\mathcal{L}$, a function of the field and its first derivatives[1].

$$\mathcal{L}(x) = \mathcal{L}\big(\phi_a(x), \partial_\mu \phi_a(x)\big) \tag{31}$$

The Principle of Least Action from Classical Mechanics states that the dynamics of a system obeying the physics described by some Lagrangian is such that the action functional is minimized. The **action** functional is defined by

$$S[\phi_a, \partial_\mu \phi_a] \equiv \int d^4x \, \mathcal{L}\big(\phi_a(x), \partial_\mu \phi_a(x)\big) \tag{32}$$

Let us consider variation by perturbing the field, but leaving the coordinates alone.

$$\delta_o \phi_a(x) = \phi'_a(x) - \phi_a(x)$$

Using the properties of variation that we have derived, we have

$$\begin{aligned}
\delta_o S &= \int d^4x \, \delta_o \mathcal{L} \\
&= \int d^4x \left[\frac{\partial \mathcal{L}}{\partial \phi_a} \, \delta_o \phi_a + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \, \underbrace{\delta_o(\partial_\mu \phi_a)}_{\partial_\mu(\delta_o \phi_a)} \right] \\
&= \int d^4x \left[\frac{\partial \mathcal{L}}{\partial \phi_a} \, \delta_o \phi_a - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \right) \delta_o \phi_a + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \, \delta_o \phi_a \right) \right] \\
&= \int d^4x \left[\frac{\partial \mathcal{L}}{\partial \phi_a} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \right) \right] \delta_o \phi_a + \underbrace{\oint d\sigma_\mu \, \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \, \delta_o \phi_a}_{0}
\end{aligned}$$

Note that terms with repeated a have an implied sum over a. We assume that we know the boundary conditions of the field, and that we are trying to derive an equation of motion for

[1] It can be shown that $\mathcal{L}$ cannot depend on higher derivatives of the field if the theory is to remain causally consistent.

the field within the boundary. Therefore, the surface integral above is zero because $\delta_o\phi_a = 0$ because we are not varying the field on the boundary.

The action is minimized at a critical point.

$$\delta_o S = 0$$

Minimizing the action for arbitrary $\delta_o\phi_a$ gives the **Euler-Lagrange Equation**.

$$0 = \int d^4x \left[\frac{\partial\mathcal{L}}{\partial\phi_a} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\right)\right]\delta_o\phi_a$$

$$\Rightarrow \qquad \boxed{\frac{\partial\mathcal{L}}{\partial\phi_a} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\right) = 0} \tag{33}$$

Given a Lagrangian, the Euler-Lagrange Equation can be used to derive the equation of motion for the field.

2.4 Noether's Theorem

Now let us allow for variations where the coordinates transform infinitesimally, called a **diffeomorphism**.

$$\delta x^\mu = x'^\mu - x^\mu$$

Then the Lagrangian varies like any other function of the coordinates.

$$\begin{aligned}
\delta\mathcal{L} &= \delta_o\mathcal{L} + \delta x^\mu\,\partial_\mu\mathcal{L} \\
&= \frac{\partial\mathcal{L}}{\partial\phi_a}\,\delta_o\phi_a + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\,\delta_o(\partial_\mu\phi_a) + \delta x^\mu\,\partial_\mu\mathcal{L} \\
&= \underbrace{\left[\frac{\partial\mathcal{L}}{\partial\phi_a} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\right)\right]}_{0}\delta_o\phi_a + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\,\delta_o\phi_a\right) + \delta x^\mu\,\partial_\mu\mathcal{L}
\end{aligned}$$

The above term is zero because $\phi_a(x)$ that is the physical solution satisfies the Euler-Lagrange Equation. Therefore, the change in a Lagrangian under a general diffeomorphism is given by the following.

$$\delta\mathcal{L} = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\,\delta_o\phi_a\right) + \delta x^\mu\,\partial_\mu\mathcal{L} \tag{34}$$

Let us consider only those diffeomorphisms that are **symmetries** of physics. That is, those transformations that leave the equations of motion invariant. This is guaranteed only if the diffeomorphism leaves the action invariant. Note that because a diffeomorphism is a change in coordinates, in general, the volume element, d^4x, also varies.

$$\delta S = \int \Big(\delta\big(d^4x\big)\mathcal{L} + d^4x\,\delta\mathcal{L}\Big) \tag{35}$$

The change in the volume element is given by the following.

$$\delta\big(d^4x\big) = (\partial_\mu \delta x^\mu)\, d^4x \tag{36}$$

Therefore

$$\begin{aligned}
\delta S &= \int d^4x \left[(\partial_\mu \delta x^\mu)\,\mathcal{L} + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\,\delta_o \phi_a\right) + \delta x^\mu\,\partial_\mu \mathcal{L}\right] \\
&= \int d^4x\,\partial_\mu \left[\mathcal{L}\,\delta x^\mu + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\,\delta_o \phi_a\right] \\
&\quad \text{using equation 27, } \delta_o\phi_a = \delta\phi_a - \delta x^\mu\,\partial_\mu \phi_a \\
&= \int d^4x\,\partial_\mu \left[\mathcal{L}\,\delta x^\mu + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\,(\delta\phi_a - \delta x^\mu\,\partial_\mu\phi_a)\right] \\
&= \int d^4x\,\partial_\mu \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\,\delta\phi_a + \left(\mathcal{L}\,\eta^{\mu\nu} - \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\,\partial^\nu \phi_a\right)\delta x_\nu\right]
\end{aligned}$$

Let us define the **energy-momentum tensor**, whose convenience will become apparent, as

$$\boxed{T^{\mu\nu} \equiv \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\,\partial^\nu \phi_a - \mathcal{L}\,\eta^{\mu\nu}} \tag{37}$$

Then

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

Now define the **Noether current** as

$$\boxed{\mathcal{J}^\mu \equiv \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)}\,\delta\phi_a - T^{\mu\nu}\,\delta x_\nu} \tag{38}$$

Because physical symmetries leave the action invariant for any region of spacetime, we can derive a conservation law as follows.

$$\delta S = \int d^4x\,\partial_\mu \mathcal{J}^\mu = 0$$

$$\Rightarrow \qquad \boxed{\partial_\mu \mathcal{J}^\mu = 0} \tag{39}$$

Therefore, for each diffeomorphism that is a symmetry of physics, there exist a conserved Noether current.

Expanding the Einstein sum in the conservation law gives

$$\partial_\mu \mathcal{J}^\mu = \partial_0 \mathcal{J}^0 - \vec{\nabla} \cdot \vec{\mathcal{J}} = 0$$

We define the **Noether charge** as

$$\boxed{\mathcal{Q} \equiv \int d^3x \, \mathcal{J}^0} \tag{40}$$

Consider the time derivative of the Noether charge. Using Stokes' Theorem, this can be converted into a surface integral. For large enough regions, on physical grounds, we expect the flux of the spacial Noether current to be zero across the boundary.

$$\frac{d\mathcal{Q}}{dt} = \int d^3x \, \partial_0 \mathcal{J}^0 = \int d^3x \, \vec{\nabla} \cdot \vec{\mathcal{J}} = \oint d\vec{\sigma} \cdot \vec{\mathcal{J}} = 0$$

Therefore the Noether charges are conserved in time.

$$\boxed{\frac{d\mathcal{Q}}{dt} = 0} \tag{41}$$

They are a formal expression for the Classically conserved quantities in physics like energy and momentum. Noether's theorem links each of these conserved quantities to a physical symmetry.

2.5 Spacetime Translation

Consider a diffeomorphism that does not mix components of the spacetime coordinates. All it does is shift each of the coordinates by a constant, c^μ.

$$x'^\mu = x^\mu + c^\mu \tag{42}$$

$$\Rightarrow \qquad \delta x^\mu = c^\mu \tag{43}$$

After the transformation, the value of the field is the same for the same spacetime point.

$$\delta\phi_a = 0 \tag{44}$$

Then, the definition of the Noether current, equation 38, gives

$$\mathcal{J}^\mu = -\, T^{\mu\nu} \, \delta x_\nu = -\, T^{\mu\nu} \, c_\nu \tag{45}$$

Because translating spacetime coordinates by an arbitrary constant is an observed symmetry in Nature, Noether's theorem says that the corresponding Noether currents are conserved.

$$\partial_\mu \mathcal{J}^\mu = 0$$

The arbitrariness of c_ν implies

$$\partial_\mu T^{\mu\nu} = 0 \tag{46}$$

We define the **total four-momentum** of the field ϕ_a as the corresponding Noether charges.

$$P^\nu \equiv \int d^3x \, T^{0\nu} \tag{47}$$

The **four-momentum density**, $\mathcal{P}^\nu$, is defined by $T^{0\nu}$, such that

$$P^\nu \equiv \int d^3x \, \mathcal{P}^\nu \tag{48}$$

Using the definition of the energy-momentum tensor, equation 37, we have

$$\mathcal{P}^\nu = \frac{\partial \mathcal{L}}{\partial(\partial_0 \phi_a)} \, \partial^\nu \phi_a - \mathcal{L} \, \eta^{0\nu} \tag{49}$$

We define the **conjugate momentum** of a field as follows.

$$\boxed{\pi_a \equiv \frac{\partial \mathcal{L}}{\partial \dot{\phi}_a}} \tag{50}$$

Then, in this notation, we have

$$\boxed{\mathcal{P}^\nu = \pi_a \, \partial^\nu \phi_a - \mathcal{L} \, \eta^{0\nu}} \tag{51}$$

The temporal component of the four-momentum density is the **Hamiltonian density**, denoted $\mathcal{H} \equiv \mathcal{P}^0$. The temporal component of the four-momentum is the **Hamiltonian**, representing the total energy of the field.

$$\boxed{H = P^0 = \int d^3x \left(\pi_a \, \dot{\phi}_a - \mathcal{L} \right) = \int d^3x \, \mathcal{H}} \tag{52}$$

The remaining spacial components are the three components of the **total momentum** of the field.

$$\boxed{\vec{P} = -\int d^3x \, \pi_a \, \vec{\nabla} \phi_a} \tag{53}$$

Energy and momentum conservation are a consequence of the translation symmetry of spacetime.

2.6 Lorentz Transformations

Now consider a diffeomorphism that mixes components of the spacetime coordinates. We know that the Lorentz transformations do this, and that physics is invariant under such a transformation. The Lorentz transformations have a unitary representation, where ${\Lambda^\mu}_\nu$ are the components of a real antisymmetric matrix.

$$x'^\mu = e^{{\Lambda^\mu}_\nu}\, x^\nu \tag{54}$$

Therefore, for infinitesimal ${\Lambda^\mu}_\nu$, the transformation can be written as

$$x'^\mu \simeq \left(\delta^\mu_\nu + {\Lambda^\mu}_\nu\right) x^\nu \tag{55}$$

$$\Rightarrow \qquad \delta x^\mu = {\Lambda^\mu}_\nu\, x^\nu \tag{56}$$

There exists a set of **spin matrices**, $\left(\Sigma^{\alpha\beta}\right)_{ab}$, labeled by four-vector indices, α and β, but the indices of a single matrix are a and b. The spin matrix relates the transformation of the spacetime four-vector coordinates to the transformation of the components of the field under a Lorentz transformation.

$$\phi'_a(x') = \left(\delta_{ab} + \frac{1}{2}\left(\Sigma^{\alpha\beta}\right)_{ab}\,\Lambda_{\alpha\beta}\right)\phi_b(x) \tag{57}$$

Note that there is implied summation over all paired indices: b, α, and β.

$$\Rightarrow \qquad \delta\phi_a(x) = \frac{1}{2}\left(\Sigma^{\alpha\beta}\right)_{ab}\,\Lambda_{\alpha\beta}\,\phi_b(x) \tag{58}$$

Plugging in equations 56 and 58 into equation 38 for the Noether current gives

$$\mathcal{J}^\mu = \frac{1}{2}\,\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\,\left(\Sigma^{\alpha\beta}\right)_{ab}\,\Lambda_{\alpha\beta}\,\phi_b - T^{\mu\alpha}\,\Lambda_{\alpha\beta}\,x^\beta$$

Using the antisymmetry of $\Lambda_{\alpha\beta}$, we have

$$\mathcal{J}^\mu = \frac{1}{2}\,\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\,\left(\Sigma^{\alpha\beta}\right)_{ab}\,\Lambda_{\alpha\beta}\,\phi_b - \frac{1}{2}\left(T^{\mu\alpha}\,x^\beta - T^{\mu\beta}\,x^\alpha\right)\Lambda_{\alpha\beta}$$

Now factoring out the arbitrary infinitesimal $\Lambda_{\alpha\beta}$ and the factor of $\frac{1}{2}$ gives

$$\mathcal{J}^\mu = \frac{1}{2}\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\,\left(\Sigma^{\alpha\beta}\right)_{ab}\,\phi_b - \left(T^{\mu\alpha}\,x^\beta - T^{\mu\beta}\,x^\alpha\right)\right]\Lambda_{\alpha\beta}$$

Let

$$M^{\mu\alpha\beta} \equiv \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\,\left(\Sigma^{\alpha\beta}\right)_{ab}\,\phi_b - \left(T^{\mu\alpha}\,x^\beta - T^{\mu\beta}\,x^\alpha\right) \tag{59}$$

Then

$$\mathcal{J}^\mu = \frac{1}{2}\, M^{\mu\alpha\beta}\, \Lambda_{\alpha\beta} \tag{60}$$

Because the Lorentz transformations are an observed symmetry in Nature, Noether's theorem says that $\partial_\mu \mathcal{J}^\mu = 0$. The arbitrariness of $\Lambda_{\alpha\beta}$ means that there is a conserved current for each α and β.

$$\partial_\mu M^{\mu\alpha\beta} = 0 \tag{61}$$

The corresponding Noether charges are then given by

$$M^{\alpha\beta} \equiv \int d^3x\, M^{0\alpha\beta} \tag{62}$$

Plugging in equations 51, 50, and 59 gives

$$\boxed{M^{\alpha\beta} = \int d^3x \left[\pi_a\, \left(\Sigma^{\alpha\beta}\right)_{ab}\, \phi_b + \left(x^\alpha\, \mathcal{P}^\beta - x^\beta\, \mathcal{P}^\alpha\right)\right]} \tag{63}$$

Like all Noether charges, each of these are conserved in time. The asymmetry of $\Lambda_{\alpha\beta}$ implies that both $\Sigma^{\alpha\beta}$ and $M^{\alpha\beta}$ are asymmetric as well.

Upon multiplication by $\frac{1}{2}$ and a Levi-Civita symbol, the spacial terms with the four-momentum density would give a cross product, $\vec{x} \times \vec{\mathcal{P}}$. This motivates us to interpret them as the orbital angular momentum. With this in mind, notice that the Σ term mixes the components of the field, a transformation in a space internal to the field as opposed to spacetime. This leads one to recognize this term as a component of angular momentum that is internal to the field, which will later be interpreted as spin of particles when the field is quantized.

$$M^{\alpha\beta} = \int d^3x \Big[\underbrace{\pi_a\, \left(\Sigma^{\alpha\beta}\right)_{ab}\, \phi_b}_{\text{spin}} + \underbrace{\left(x^\alpha\, \mathcal{P}^\beta - x^\beta\, \mathcal{P}^\alpha\right)}_{\text{orbital angular momentum}}\Big] \tag{64}$$

The angular momentum is given by

$$J^k = \frac{1}{2}\epsilon^{ijk} M_{ij} \tag{65}$$

Angular momentum conservation is a consequence of the rotation symmetry of spacetime, part of Lorentz invariance. The M^{0i} charges, derived from the transformations mixing space and time components, correspond to Lorentz boosts. They too are conserved, but are not as physically apparent as angular momentum.

2.7 Internal Symmetries

In addition to the symmetries of spacetime, some fields have symmetries relating their internal degrees of freedom. These types of transformations have the following unitary representation

$$\phi'_a(x) = e^{i\,(G_r)_{ab}\,\theta_r}\,\phi_b(x) \tag{66}$$

where G_r are hermitian matrices, such that the transformation is unitary. (Note that summation is implied over r, a, and b.) G_r are called the "**generators**" of the transformation, and θ_r are the corresponding parameters. Note that this type of transformation has nothing to do with spacetime. Instead, it directly mixes the components of the field. If this type of transformation leaves the equations of motion invariant, and is therefore a symmetry, then physics is unaffected by whether phenomena are described by $\phi_a(x)$, or a field related to $\phi_a(x)$ by such a transformation. Fields related to one another by such a transformation are called different "**gauges**" of the same field, and this type of symmetry is known as "**gauge invariance**."

For infinitesimal θ_r, the transformation can be expanded as

$$\phi'_a(x) \simeq \big(\delta_{ab} + i\,(G_r)_{ab}\,\theta_r\big)\phi_b(x) \tag{67}$$

$$\Rightarrow \qquad \delta\phi_a(x) = i\,(G_r)_{ab}\,\theta_r\,\phi_b(x) \tag{68}$$

Plugging this into equation 38 gives the Noether currents.

$$\mathcal{J}^\mu_r = i\,\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\,(G_r)_{ab}\,\theta_r\,\phi_b(x) \tag{69}$$

If the Lagrangian is gauge invariant, then those Noether currents are conserved, and the corresponding Noether charges (derived using equations 40 and 50) are conserved in time.

$$\boxed{Q_r = \int d^3x\; i\,\pi_a\,(G_r)_{ab}\,\theta_r\,\phi_b} \tag{70}$$

3 The Free Real Scalar Field

We will investigate the simplest example of a field, a real scalar field. That is, one that has a single component that is a real number and therefore transforms trivially under the Lorentz transformations. After studying the Classical properties of the real scalar field, we will summarize the principles of quantum mechanics, and finally quantize the real scalar field. We will use this simple example to explain the canonical quantization procedure that can then be applied to any other type of field.

3.1 The Classical Theory

The Classical theory of a field has two inputs: the type of field, and the Lagrangian that describes its dynamics. The field we are now studying is the real scalar field that we will denote $\phi(x)$. The Lagrangian[2] describing the free relativistic dynamics of this field is

$$\mathcal{L} = \frac{1}{2}\big(\,(\partial_\mu \phi)\,(\partial^\mu \phi) - m^2\,\phi^2\big) \tag{71}$$

where m is a real constant.

The first step in studying the Classical theory of a field is to calculate its equation of motion and its Noether charges. The equation of motion, calculated by plugging the Lagrangian into the Euler-Lagrange equation 33, is

$$\boxed{\left(\partial^2 + m^2\right)\phi = 0} \tag{72}$$

This is called the **Klein-Gordon equation**. Equation 50 says that the conjugate momentum for the real scalar field is

$$\boxed{\pi(x) = \dot{\phi}(x)} \tag{73}$$

The Hamiltonian and momentum are given by equations 52 and 53.

$$\boxed{H = \int d^3x\, \frac{1}{2}\left(\pi^2 + (\vec{\nabla}\phi)^2 + m^2\,\phi^2\right)} \tag{74}$$

$$\boxed{\vec{P} = -\int d^3x\, \pi\, \vec{\nabla}\phi} \tag{75}$$

This implies that the four-momentum density is given by

$$\mathcal{P}^0 = \mathcal{H} = \frac{1}{2}\left(\pi^2 + (\vec{\nabla}\phi)^2 + m^2\,\phi^2\right) \tag{76}$$

[2]The appropriateness of this Lagrangian will become apparent when we see that it insures that the field is relativistic by satisfying equation 11.

and

$$\vec{\mathcal{P}} = -\,\pi\,\vec{\nabla}\phi \tag{77}$$

which is used in equation 63 to calculate $M^{\mu\nu}$.

$$\boxed{M^{\mu\nu} = \int d^3x\,(x^\mu\,\mathcal{P}^\nu - x^\nu\,\mathcal{P}^\mu)} \tag{78}$$

Note that the spin term in $M^{\mu\nu}$ is zero because the real scalar field does not change under the Lorentz transformation and therefore equation 57 implies that the spin matrix, Σ, is zero.

We next turn our attention studying the solutions of the Klein-Gordon equation. Consider the **plane-waves** $e^{i\,k\cdot x}$ of various k. Plugging this into the Klein-Gordon equation gives

$$(\partial^2 + m^2)\,e^{i\,k\cdot x} = \left((i\,k)^2 + m^2\right)\,e^{i\,k\cdot x} = (-k^2 + m^2)\,e^{i\,k\cdot x} = 0\,.$$

$$\Rightarrow \qquad k^2 = m^2$$

This shows that $e^{i\,k\cdot x}$ is a solution if and only if we constrain one of the four degrees of freedom in the four-vector k such that $k^2 = \omega^2 - \vec{k}^2 = m^2$. We do this by setting

$$k^0 = \omega_k \equiv \sqrt{\vec{k}^2 + m^2}\,. \tag{79}$$

As one knows from Fourier analysis, the functions $\left\{e^{i\,k\cdot x}\right\}$ for various k form a complete basis for a complex function space. We can decompose any complex function $\phi(x)$ satisfying the Klein-Gordon equation by

$$\phi(x) = \int \frac{d^3k}{(2\pi)^3}\,b(k)\,\left(a(k)e^{-i\,k\cdot x} + a^*(k)e^{i\,k\cdot x}\right)\,.$$

We are integrating over the three-momentum $\vec{k}$ instead of all four spacetime components because we constrained k by equation 79 to guaranty that $\phi(x)$ is a solution of the Klein-Gordon equation. The three-momentum integral is not manifestly invariant like an integral over d^4k, so we have inserted the function $b(k)$ which we will use to constrain the rest of the integrand to be Lorentz invariant. The $a(k)$ are the coefficients of the Fourier expansion. We have added the complex conjugate to ensure that $phi(x)$ is real.

Note that three-momentum varies under Lorentz boosts, so we need to find $b(k)$ such that $d^3k\,b(k)$ is the Lorentz invariant three-momentum integration measure. $\int d^4k\,L(k)$ is manifestly invariant if $L(k)$ is a Lorentz invariant scalar. How can we make an integral over d^3k Lorentz invariant? We can take a Lorentz invariant integral over d^4k and constrain one of the degrees of freedom, while preserving the Lorentz invariance by forcing k^2 to always remain m^2 with a delta function.

$$\int \frac{d^4k}{(2\pi)^4}\, 2\pi\, \delta(k^2 - m^2)\, \theta(k^0)\, L(k)$$

The factor of 2π is just a convention because we are going to integrate over one degree

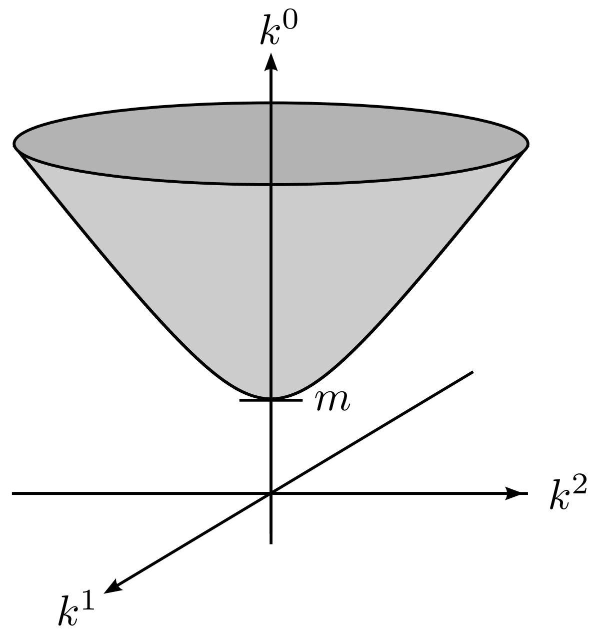


Figure 1: The hyperboloid over which Lorentz invariant three-momentum integration is performed.

of freedom and we want to be left with a factor of $1/(2\pi)$ for each remaining momentum integral. Obviously it has no effect on the Lorentz invariance of the result. The step function limits the integration to the region where $k^0 \geq 0$ because only four-vectors with non-negative energy are reasonable. Visually, this is an integral over the hyperboloid where $k^2 = m^2$, or $k^0 = \sqrt{m^2 + \vec{k}^2}$, shown in Figure 1. One can simplify this integral with the following identity for delta functions

$$\delta(g(x)) = \sum_i \frac{\delta(x - x_i)}{|g'(x_i)|}\,, \tag{80}$$

where $\{x_i\}$ are the zeros of $g(x)$ and $g'(x)$ denotes the derivative of $g(x)$. Note that

$$\frac{d}{dk^0}(k^2 - m^2) = 2k^0 = 2\omega_k\,,$$

therefore

$$
\begin{aligned}
\int \frac{d^4k}{(2\pi)^4}\, & 2\pi\, \delta(k^2 - m^2)\, \theta(k^0)\, L(k) \\
&= \int \frac{d^3k}{(2\pi)^3} \int dk^0 \left(\frac{\delta(k^0 - \omega_k)}{2\,\omega_k} + \frac{\delta(k^0 + \omega_k)}{2\,\omega_k} \right) \theta(k^0)\, L(k) \\
&= \int \frac{d^3k}{(2\pi)^3}\, \frac{1}{2\omega_k}\, L(k)\, .
\end{aligned}
$$

Therefore, $\frac{d^3k}{(2\pi)^3\, 2\omega_k}$ is the Lorentz invariant three-momentum measure, also called the Lorentz invariant phase space element.

Evidently, $b(k) = 1/(2\omega_k)$ and our Fourier decomposition of $\phi(x)$ should be written as

$$
\boxed{\phi(x) = \int \frac{d^3k}{(2\pi)^3\, 2\omega_k} \left(a(k) e^{-i\,k\cdot x} + a^*(k) e^{i\,k\cdot x} \right)}\, . \tag{81}
$$

We can project out the value $a(k)$ for a specific k, but first we need to derive a couple completeness orthogonality relations for this basis of functions $\left\{ e^{i\,k\cdot x} \right\}$. These relations use an inner product that involves the following operation, which takes the time derivative to the right first and then subtracts the time derivative acting to the left.

$$
a\, \overleftrightarrow{\partial_0}\, b \equiv a\, \partial_0 b - (\partial_0 a)\, b \tag{82}
$$

The completeness relation is derived by the following.

$$
\begin{aligned}
\int d^3x\, e^{i\,k'\cdot x}\, i\, \overleftrightarrow{\partial_0}\, e^{-i\,k\cdot x} &= i \int d^3x \left[e^{i\,k'\cdot x}\, \partial_0 e^{-i\,k\cdot x} - \left(\partial_0 e^{i\,k'\cdot x} \right) e^{-i\,k\cdot x} \right] \\
&= (\omega_k + \omega_{k'}) \int d^3x\, e^{i(k'-k)\cdot x} \\
&= (\omega_k + \omega_{k'})\, \underbrace{e^{i(\omega_{k'} - \omega_k)t}}_{e^0 = 1}\, \underbrace{\int d^3x\, e^{-i(\vec{k}' - \vec{k})\cdot \vec{x}}}_{(2\pi)^3\, \delta^3\left(\vec{k} - \vec{k}'\right)} \\
&= 2\omega_k\, (2\pi)^3\, \delta^3\left(\vec{k} - \vec{k}'\right)\, .
\end{aligned} \tag{83}
$$

Similarly, the orthogonality relation is derived by

$$\begin{aligned}
\int d^3x\, e^{i\,k'\cdot x}\, i\, \overleftrightarrow{\partial_0}\, e^{i\,k\cdot x} &= i \int d^3x \left[e^{i\,k'\cdot x}\, \partial_0 e^{i\,k\cdot x} - \left(\partial_0 e^{i\,k'\cdot x}\right) e^{i\,k\cdot x} \right] \\
&= (\omega_k - \omega_{k'}) \int d^3x\, e^{i(k'+k)\cdot x} \\
&= \cancelto{0}{(\omega_k - \omega_{k'})}\, e^{i(\omega_{k'}+\omega_k)t} \underbrace{\int d^3x\, e^{-i(\vec{k}'+\vec{k})\cdot\vec{x}}}_{(2\pi)^3\, \delta^3\left(\vec{k}+\vec{k}'\right)} \\
&= 0\,. \qquad (84)
\end{aligned}$$

Using the completeness and orthogonality relations, we can invert this expansion to find an expression for the expansion coefficients $a(k)$.

$$\begin{aligned}
\int d^3x\, e^{i\,k\cdot x}\, i\, \overleftrightarrow{\partial_0}\, \phi(x) = \int \frac{d^3k}{(2\pi)^3} \Big[a(k) & \underbrace{\int d^3x\, e^{i\,k\cdot x}\, i\, \overleftrightarrow{\partial_0}\, e^{-i\,k\cdot x}}_{(2\pi)^3\, \delta^3\left(\vec{k}-\vec{k}\right)} \\
+\; a^*(k) & \underbrace{\int d^3x\, e^{i\,k\cdot x}\, i\, \overleftrightarrow{\partial_0}\, e^{i\,k\cdot x}}_{0} \Big]
\end{aligned}$$

$$\therefore \qquad \boxed{a(k) = i \int d^3x\, e^{i\,k\cdot x}\;\; \overleftrightarrow{\partial_0}\, \phi(x)}\,. \qquad (85)$$

We can write this in terms of the conjugate momentum by letting the $\overleftrightarrow{\partial_0}$ act on ϕ.

$$\begin{aligned}
a(k) &= i \int d^3x \big[e^{i\,k\cdot x}\, \partial_0\phi(x) - \phi(x)\, \partial_0 e^{i\,\cdot x} \big] \\
&= i \int d^3x \big[e^{i\,k\cdot x}\, \pi(x) - \phi(x)\, (i\,\omega_k)\, e^{i\,k\cdot x} \big]
\end{aligned}$$

$$\therefore \qquad \boxed{a(k) = \int d^3x\, e^{i\,k\cdot x} \left(\omega_k\, \phi(x) + i\,\pi(x) \right)} \qquad (86)$$

3.2 Principles of Quantum Mechanics

It is now appropriate to pause our study of the real scalar field and to outline the canonical quantization procedure for how one goes from studying a classical field to a quantum field in general. First we will summarize what **quantum mechanics** means.

3.2.1 States

Quantum mechanics fundamentally concerns the study of the dynamics of **states** and **operators** in a **Hilbert space**. A Hilbert space is a complex vector space that often has an infinite number dimensions, where a vector, $|\alpha\rangle$, denotes the state of a quantum system. In general, a vector can be expanded in terms of a complete orthonormal basis $\{|b_i\rangle\}$.

$$|\alpha\rangle = \sum_i a_i \; |b_i\rangle$$

In the case that we use a basis that is parametrized by a real number, x, instead of integers, the sum becomes an integral.

$$|\alpha\rangle = \int dx \, a(x) \; |x\rangle$$

There exists a **dual** vector space that is the adjoint of the first.

$$\langle\alpha| \equiv |\alpha\rangle^\dagger = \sum_i a_i^* \; \langle b_i|$$

A Hilbert space has an binary operation between a vector and a dual vector called an **inner product**. The inner product of two vectors $|\alpha\rangle$ and $|\beta\rangle$ denoted as $\langle\alpha|\beta\rangle$. Two vectors are **orthogonal** if there inner product is zero.

$$\langle\alpha|\beta\rangle = 0$$

A set of vectors, $\{|b_i\rangle\}$, is an **orthonormal basis** if every vector is orthogonal with every other vector and the inner product of every vector with itself is one.

$$\langle b_i|b_j\rangle = \delta_{ij}$$

The basis is **complete** if they span the space and therefore the following sum is unity.

$$\sum_i |b_i\rangle \langle b_i| = \hat{\mathbb{1}}$$

Using an orthonormal basis, one can decompose any vector into a sum over the basis vectors, whose coefficients are given by the inner products of the basis vectors and the vector being decomposed.

$$|\alpha\rangle = \sum_i |b_i\rangle \underbrace{\langle b_i|\alpha\rangle}_{a_i} = \sum_i a_i \; |b_i\rangle$$

We can see why we choose the dual vectors be the adjoint of vectors because it conveniently forces the inner product of a vector with itself, its **norm**, to be real and nonnegative.

$$\begin{aligned}
\langle\alpha|\alpha\rangle &= \left(\sum_i a_i^* \langle b_i|\right)\left(\sum_j a_j |b_j\rangle\right) \\
&= \sum_i a_i^*\, a_i \cancelto{1}{\langle b_i|b_i\rangle} \\
&= \sum_i a_i^*\, a_i \geq 0
\end{aligned}$$

This is because $a_i^*\, a_i$ is real and nonnegative for any complex number a_i. We normalize state vectors by constraining the above sum to be one.

$$\langle\alpha|\alpha\rangle = 1$$

3.2.2 Operators

Like any vector space, one can define linear operators, denoted with hâts, to act on a vector and return another vector.

$$\hat{O}\,|\psi\rangle = |\psi'\rangle$$

Linear means that

$$\hat{O}\big(c_1\,|\alpha\rangle + c_2\,|\beta\rangle\,\big) = c_1\,\hat{O}\,|\alpha\rangle + c_2\,\hat{O}\,|\beta\rangle$$

where c_1 and c_2 are some complex numbers.

Physical quantities like four-momentum, charge, etc. are represented by operators. When a state is an eigenstate of an operator, we interpret it as having a well defined value for that physical quantity.

$$\hat{B}_i\ |b_i\rangle = b_i\ |b_i\rangle$$

In the above equation, the state $|b_i\rangle$ is interpreted as being a state with a well defined value b_i for the physical quantity represented by the operator $\hat{B}_i$. Operators representing physical quantities must always be hermitian such that they real eigenvalues.

If one chooses operators that have eigenstates that form a complete orthonormal basis, one can decompose a general vector into a sum over eigenstates.

$$|\psi\rangle = \sum_i a_i\ |b_i\rangle$$

Such a state with nonzero values for more than one a_i is said to be in a quantum mechanical *superposition* of the corresponding eigenstates. For example, the following state has no well defined value for the quantity represented by $\hat{B}$.

$$|\psi\rangle = \frac{1}{\sqrt{2}}\big(\,|b_1\rangle + |b_2\rangle\,\big)$$

Instead, it has some probability to have the value b_1 and some probability to have the value b_2. The probabilities are determined by calculating the magnitude squared of the coefficients in the expansion. In this case, $(1/\sqrt{2})^2 = 0.5$, so each value has a 50% probability. In general, an inner product can be thought of as selecting out the coefficient corresponding to a basis vector in the expansion of some vector.

$$\langle b_j|\psi\rangle = \langle b_j|\,\sum_i a_i\,\,|b_i\rangle = \sum_i a_i\,\,\langle b_j|b_i\rangle = a_j$$

An inner product of a state vector with an eigenvector for some observable is known as the quantum mechanical **amplitude** for the state to have that value for that observable, and the probability for it is the magnitude squared of the amplitude, $|\langle b_j|\psi\rangle|^2$. An amplitude, $\langle\beta|\psi\rangle$, is also referred to as the quantum mechanical *overlap* between the states $|\beta\rangle$ and $|\psi\rangle$.

3.2.3 The Fundamental Postulate of Quantum Mech.

When one quantizes a classical field theory, the field $\phi(x)$, and its canonical momentum $\pi(x)$ become the operators, $\hat{\phi}(x)$ and $\hat{\pi}(x)$ in a Hilbert space. Classically, $\phi(x)$ and $\pi(x)$, are determined by the initial conditions of the field (like initial position and momentum for point particle mechanics) and contain all the information of the system. Similarly, all operators in the Hilbert space can be written in terms of the operators $\hat{\phi}(x)$ and $\hat{\pi}(x)$. For example, we will see that the Noether's charges P^μ and $M^{\mu\nu}$ become the operators $\hat{P}^\mu$ and $\hat{M}^{\mu\nu}$, written in terms of $\hat{\phi}(x)$ and $\hat{\pi}(x)$. This will be explained more concretely when we discuss the specific example of the quantum theory of the real scalar field.

Since we are going to represent the state of a physical system by a vector in the Hilbert space, we need to understand how state vectors and operators change when the physical system changes. How do we represent transformations like spacetime translation and Lorentz transformations in the Hilbert space? For each type of transformation, there must be some operator, $\hat{U}(\theta_a)$, that acts on a state vector, $|\alpha\rangle$, and returns a state vector, $|\alpha'\rangle$, corresponding to the state transformed by some amount parametrized by the parameters θ_a.

$$|\alpha'\rangle = \hat{U}(\theta_a)\,|\alpha\rangle \tag{87}$$

Because we want to keep state vectors normalized after such an operation, the operators representing transformations must be unitary.

$$\langle\alpha'|\alpha'\rangle = \langle\alpha|\,\hat{U}^\dagger(\theta_a)\,\hat{U}(\theta_a)\,|\alpha\rangle = \langle\alpha|\alpha\rangle = 1$$

$$\Rightarrow \qquad \hat{U}^{\dagger}(\theta_a)\, \hat{U}(\theta_a) = \hat{\mathbb{1}}$$

Any unitary operator can be written as

$$\hat{U}(\theta_a) = e^{-i\, \hat{G}_a\, \theta_a}$$

where $\hat{G}_a$ are hermitian, insuring that $\hat{U}(\theta_a)$ is unitary. In the limit θ_a go to zero, the transformation operator becomes the identity, as one would expect.

$$\hat{U}(\theta_a) = \hat{\mathbb{1}} - i\, \hat{G}_a\, \theta_a + \dots$$

Similar to the generators discussed in the Section 2.7 on internal symmetries, $\hat{G}_a$ are know as the "generators" of some transformation.

We represent spacetime translations in the Hilbert space by

$$\hat{U}(x^{\mu}) = e^{-i\, \hat{T}_{\mu}\, x^{\mu}}$$

where x^{μ} the spacetime four-vector by which the system was translated. And we represent the Lorentz transformations in the Hilbert space by

$$\hat{U}(\theta^{\mu\nu}) = e^{-i\, \hat{L}_{\mu\nu}\, \theta^{\mu\nu}}$$

What could these generators be? What quantities are fundamentally associated with the spacetime translation and the Lorentz transformations? Noether's theorem points out that the Noether charges are fundamentally linked to the symmetry of a transformation. Following this hint, we state what this text calls **The Fundamental Postulate of Quantum Mechanics:**[3]

> The generators of the representation of a transformation in the Hilbert space are the operators representing the classical Noether's Charges that are conserved under that transformation.

In other words, $\hat{T}_{\mu} = \hat{P}_{\mu}$ and $\hat{L}_{\mu\nu} = \frac{1}{2}\hat{M}_{\mu\nu}$ (the $\frac{1}{2}$ is just a convention). Therefore, the operation of translating a system by x^{μ} in spacetime is represented by the following in the Hilbert space.

$$\hat{U}(x^{\mu}) = e^{-i\, \hat{P}_{\mu}\, x^{\mu}} \tag{88}$$

Similarly, a Lorentz transformation where $\theta^{\mu\nu}$ parametrizes a combination of boosts and rotations, is represented by

$$\hat{U}(\theta^{\mu\nu}) = e^{-i\, \frac{1}{2}\, \hat{M}_{\mu\nu}\, \theta^{\mu\nu}} \tag{89}$$

[3] TODO: Note this is actually a rephrasing of Wigner's theorem as the cornerstone of quantum mechanics.

in the Hilbert space. You may have learned in a quantum mechanics class that the Hamiltonian is the generator of time translation.

$$\hat{U}(t) = e^{-i\,\hat{H}\,t}$$

Equations 88 and 89 are a generalization of that fact.

In equation 87, we introduced the representations of transformations as fixed operators that act on and transform state vectors. This is called the "Schrödinger picture." But there is an alternate picture, called the "Heisenberg picture," where the state vectors are fixed and the operators are dynamic. We will more often use the Heisenberg picture, where the operators $\hat{\phi}(x)$ and $\hat{\pi}(x)$ are dynamic in the quantum theory like the fields $\phi(x)$ and $\pi(x)$ were dynamic in the Classical theory. From now on in this text, we will use Heisenberg Operators exclusively unless otherwise noted. What is important is that the amplitudes and expectation values are the same in either picture the Heisenberg or Schrödinger picture.

Consider a state translated in spacetime by ϵ^μ.

$$|\alpha+\epsilon\rangle = e^{-i\,\hat{P}_\mu\,\epsilon^\mu}\,|\alpha\rangle$$

The expectation value of some operator, $\hat{O}(x)$, for this state is

$$\langle\alpha+\epsilon|\hat{O}(x)|\alpha+\epsilon\rangle = \langle\alpha|\,e^{i\,\hat{P}_\mu\,\epsilon^\mu}\,\hat{O}(x)\,e^{-i\,\hat{P}_\mu\,\epsilon^\mu}\,|\alpha\rangle$$

Here we see that as we translate the system by ϵ^μ, we can either think of transforming the vector or the operator. Evidently,

$$\hat{O}(x+\epsilon) = e^{i\,\hat{P}_\mu\,\epsilon^\mu}\,\hat{O}(x)\,e^{-i\,\hat{P}_\mu\,\epsilon^\mu} \tag{90}$$

is how Heisenberg Operators[4] transform under translations in spacetime. Consider an infinitesimal transformation.

$$\begin{aligned}
\hat{O}(x) + \epsilon^\mu\,\partial_\mu\hat{O}(x) &= (1+i\,\hat{P}_\mu\,\epsilon^\mu)\hat{O}(x)(1-i\,\hat{P}_\mu\,\epsilon^\mu) + \mathcal{O}[\epsilon^2] \\
&= \hat{O}(x) + i\,\hat{P}_\mu\,\hat{O}(x)\,\epsilon^\mu - i\,\hat{P}_\mu\,\hat{O}(x)\,\epsilon^\mu + \mathcal{O}[\epsilon^2] \\
&= \hat{O}(x) + i\,[\hat{P}_\mu,\hat{O}(x)]\,\epsilon^\mu
\end{aligned}$$

$$\therefore \qquad \boxed{\partial_\mu\hat{O}(x) = i[\hat{P}_\mu,\hat{O}(x)]} \tag{91}$$

This is the **Heisenberg Equation of Motion**. It determines the equations of motion for operators much like the Euler-Lagrange equation determines the equations of motion for Classical fields.

[4] operators in the Heisenberg Picture

3.2.4 The Poincaré Algebra

Spacetime is observed to have certain symmetries, namely spacetime translation invariance, and Lorentz invariance, including boosts and rotations. Together, the group of these transformations is collectively called the **Poincaré group**. If we are going to represent physical systems in spacetime by states in a Hilbert space, then those states should observe those same symmetries. This means that the generators of those transformations in the Hilbert space must satisfy the **Poincaré algebra**:

$$\left[\hat{P}_\mu, \hat{P}_\nu\right] = 0 \tag{92}$$

$$\left[\hat{M}_{\mu\nu}, \hat{P}_\lambda\right] = -i\,\eta_{\mu\lambda}\,\hat{P}_\nu + i\,\eta_{\nu\lambda}\,\hat{P}_\mu \tag{93}$$

$$\left[\hat{M}_{\mu\nu}, \hat{M}_{\lambda\sigma}\right] = i\,\eta_{\nu\lambda}\,\hat{M}_{\mu\sigma} - i\,\eta_{\mu\lambda}\,\hat{M}_{\nu\sigma} - i\,\eta_{\nu\sigma}\,\hat{M}_{\mu\lambda} + i\,\eta_{\mu\sigma}\,\hat{M}_{\nu\lambda} \tag{94}$$

Plugging in $\hat{P}_\nu$ for $\hat{O}$ in the Heisenberg Equation of Motion, and using the Poincaré algebra, shows that the operator $\hat{P}_\nu$ is conserved just like its classical Noether's charge. Similarly, $\hat{M}_{\mu\nu}$ is conserved.

$$\partial_0 \hat{P}_\nu = 0 \qquad\qquad \partial_0 \hat{M}_{\mu\nu} = 0$$

3.2.5 Canonical Quantization

We will now outline the procedure for **canonical quantization**, a method for turning a classical theory into a quantum theory by representing it in a Hilbert space.

First, the dynamical variables of the classical theory become operators in the Hilbert space. For a field theory, the dynamical variables are the field, $\phi(x)$, and its conjugate momentum, $\pi(x)$. The classical Noether's charges become operators as well. The operators $\hat{P}^\mu$ and $\hat{M}^{\mu\nu}$ are defined in terms of the same classical relations as the functions P^μ and $M^{\mu\nu}$, with $\phi(x)$ and $\pi(x)$ changed to their corresponding operators. Following the fundamental postulate of quantum mechanics, we choose the operators representing the corresponding Noether charges to generate the Poincaré transformations. Therefore we must *demand* that those operators satisfy the Poincaré algebra.

One should now ask what effect the Poincaré algebra has on the field operators $\hat{\phi}(x)$ and $\hat{\pi}(x)$. We will see when we study the quantum theory of the real scaler field as an example, that demanding that $\hat{P}^\mu$ and $\hat{M}^{\mu\nu}$ satisfy the Poincaré algebra *imposes* **canonical commutation relations** on $\hat{\phi}(x)$ and $\hat{\pi}(x)$. These commutation relations are what generate all the hallmarks quantum mechanics including uncertainty relations and the spin-statistics theorem.

To clarify the details of this procedure, we now turn our attention back to the real scalar field to quantize it.

3.3 The Quantum Theory

The dynamical variables $\hat{\phi}(x)$ and $\hat{\pi}(x)$ are now operators in a Hilbert space. Equations 74, 75, and 78 with the dynamical variables turned into operators, give expressions for the operators $\hat{P}^\mu$ and $\hat{M}^{\mu\nu}$. We will now use these expressions and the Poincaré algebra to derive the canonical commutation relations, also known as the "**equal time commutation relations**."

3.3.1 Equal Time Commutation Relations

Equation 75 implies that

$$\hat{P}_j(x^0) = \int d^3x\, \hat{\pi}(x)\, \partial_j \hat{\phi}(x) \tag{95}$$

We require that this satisfy equation 92 from the Poincaré algebra.

$$\left[\hat{P}_j(x^0), \hat{P}_k(y^0)\right] = 0$$

Evaluating this commutator in terms of $\hat{\phi}(x)$ and $\hat{\pi}(x)$, we have

$$\begin{aligned} \left[\hat{P}_j(x^0), \hat{P}_k(y^0)\right] &= \int d^3x\, d^3y\, \left[\hat{\pi}(x)\, \partial_{x^j} \hat{\phi}(x), \hat{\pi}(y)\, \partial_{y^k} \hat{\phi}(y)\right] \\ &= \int d^3x\, d^3y \Big(\hat{\pi}(x)\, \partial_{x^j} \hat{\phi}(x)\, \hat{\pi}(y)\, \partial_{y^k} \hat{\phi}(y) \\ &\quad - \hat{\pi}(y)\, \partial_{y^k} \underbrace{\hat{\phi}(y)\, \hat{\pi}(x)}_{\text{commute}}\, \partial_{x^j} \hat{\phi}(x)\Big) \end{aligned}$$

We will commute $\hat{\phi}$ and $\hat{\pi}$ several times in this derivation asuming they have some commutation relation. Note that for any operators $\hat{A}$ and $\hat{B}$,

$$\hat{A}\, \hat{B} = \left[\hat{A}, \hat{B}\right]_{\pm} \mp \hat{B}\, \hat{A}$$

where $\left[\hat{A}, \hat{B}\right]_{-}$ is a commutator and $\left[\hat{A}, \hat{B}\right]_{+}$ is an anticommutator.

$$\left[\hat{A}, \hat{B}\right]_{-} \equiv \left[\hat{A}, \hat{B}\right] \equiv \hat{A}\, \hat{B} - \hat{B}\, \hat{A}$$

$$\left[\hat{A}, \hat{B}\right]_{+} \equiv \left\{\hat{A}, \hat{B}\right\} \equiv \hat{A}\, \hat{B} + \hat{B}\, \hat{A}$$

Using this to successively commute operators, we get

$$
\begin{aligned}
& \left[\hat{P}_j(x^0), \hat{P}_k(y^0)\right] \\
= & \int d^3x\, d^3y \Big(\hat{\pi}(x)\, \partial_{x^j} \hat{\phi}(x)\, \hat{\pi}(y)\, \partial_{y^k} \hat{\phi}(y) - \hat{\pi}(y)\, \partial_{y^k} \left[\hat{\phi}(y), \hat{\pi}(x)\right]_\pm \partial_{x^j} \hat{\phi}(x) \\
& \pm \underbrace{\hat{\pi}(y)\, \hat{\pi}(x)}_{\text{commute}} \partial_{y^k} \hat{\phi}(y)\, \partial_{x^j} \hat{\phi}(x) \Big) \\
= & \int d^3x\, d^3y \Big(\hat{\pi}(x)\, \partial_{x^j} \hat{\phi}(x)\, \hat{\pi}(y)\, \partial_{y^k} \hat{\phi}(y) - \hat{\pi}(y)\, \partial_{y^k} \left[\hat{\phi}(y), \hat{\pi}(x)\right]_\pm \partial_{x^j} \hat{\phi}(x) \\
& \pm \left[\hat{\pi}(y), \hat{\pi}(x)\right]_\pm \partial_{y^k} \hat{\phi}(y)\, \partial_{x^j} \hat{\phi}(x) - \hat{\pi}(x)\, \hat{\pi}(y)\, \underbrace{\partial_{y^k} \hat{\phi}(y)\, \partial_{x^j} \hat{\phi}(x)}_{\text{commute}} \Big) \\
= & \int d^3x\, d^3y \Big(\hat{\pi}(x)\, \partial_{x^j} \hat{\phi}(x)\, \hat{\pi}(y)\, \partial_{y^k} \hat{\phi}(y) - \hat{\pi}(y)\, \partial_{y^k} \left[\hat{\phi}(y), \hat{\pi}(x)\right]_\pm \partial_{x^j} \hat{\phi}(x) \\
& \pm \left[\hat{\pi}(y), \hat{\pi}(x)\right]_\pm \partial_{y^k} \hat{\phi}(y)\, \partial_{x^j} \hat{\phi}(x) - \hat{\pi}(x)\, \hat{\pi}(y)\, \partial_{y^k} \partial_{x^j} \left[\hat{\phi}(y), \hat{\phi}(x)\right]_\pm \\
& \pm \hat{\pi}(x)\, \underbrace{\hat{\pi}(y)\, \partial_{x^j} \hat{\phi}(x)}_{\text{commute}}\, \partial_{y^k} \hat{\phi}(y) \Big) \\
= & \int d^3x\, d^3y \Big(\cancel{\hat{\pi}(x)\, \partial_{x^j} \hat{\phi}(x)\, \hat{\pi}(y)\, \partial_{y^k} \hat{\phi}(y)} - \overbrace{\hat{\pi}(y)\, \partial_{y^k} \left[\hat{\phi}(y), \hat{\pi}(x)\right]_\pm \partial_{x^j} \hat{\phi}(x)}^{x \leftrightarrow y} \\
& \pm \left[\hat{\pi}(y), \hat{\pi}(x)\right]_\pm \partial_{y^k} \hat{\phi}(y)\, \partial_{x^j} \hat{\phi}(x) - \hat{\pi}(x)\, \hat{\pi}(y)\, \partial_{y^k} \partial_{x^j} \left[\hat{\phi}(y), \hat{\phi}(x)\right]_\pm \\
& \pm \hat{\pi}(x)\, \partial_{x^j} \left[\hat{\pi}(y), \hat{\phi}(x)\right]_\pm \partial_{y^k} \hat{\phi}(y) - \cancel{\hat{\pi}(x)\, \partial_{x^j} \hat{\phi}(x)\, \hat{\pi}(y)\, \partial_{y^k} \hat{\phi}(y)} \Big)
\end{aligned}
$$

Note that in the second term, we are allowed to interchange the four-vectors x and y with no effect on $\left[\hat{P}_j(x^0), \hat{P}_k(y^0)\right]$ only if it is an equal time commutator, $x^0 = y^0$. The three-vector components of x and y are irrelevant because they are integrated over. Therefore all the commutators in this derivation are evaluated at *equal times*. In that same term, we also flip the commutator.

$$
\begin{aligned}
& \left[\hat{P}_j, \hat{P}_k\right] \\
= & \int d^3x\, d^3y \Big(\mp \hat{\pi}(y)\, \partial_{y^k} \left[\hat{\pi}(y), \hat{\phi}(x)\right]_\pm \partial_{x^j} \hat{\phi}(x) \pm \left[\hat{\pi}(y), \hat{\pi}(x)\right]_\pm \partial_{y^k} \hat{\phi}(y)\, \partial_{x^j} \hat{\phi}(x) \\
& - \hat{\pi}(x)\, \hat{\pi}(y)\, \partial_{y^k} \partial_{x^j} \left[\hat{\phi}(y), \hat{\phi}(x)\right]_\pm \pm \hat{\pi}(x)\, \partial_{x^j} \left[\hat{\pi}(y), \hat{\phi}(x)\right]_\pm \partial_{y^k} \hat{\phi}(y) \Big)
\end{aligned}
$$

Now we integrate by parts to get the derivatives off of the commutators. Note that all the surface terms are zero because the field dies off at infinity. Watch the signs.

$$\begin{aligned}&[\hat{P}_j, \hat{P}_k] \\ =\ & \int d^3x\, d^3y \Big(\pm \big(\partial_{y^k}\hat{\pi}(y)\big)\, \big[\hat{\pi}(y), \hat{\phi}(x)\big]_\pm\, \partial_{x^j}\hat{\phi}(x)\ \pm\ \big[\hat{\pi}(y), \hat{\pi}(x)\big]_\pm\, \partial_{y^k}\hat{\phi}(y)\, \partial_{x^j}\hat{\phi}(x) \\ & - \big(\partial_{x^j}\hat{\pi}(x)\big)\, \big(\partial_{y^k}\hat{\pi}(y)\big)\, \big[\hat{\phi}(y), \hat{\phi}(x)\big]_\pm\ \mp\ \big(\partial_{x^j}\hat{\pi}(x)\big)\, \big[\hat{\pi}(y), \hat{\phi}(x)\big]_\pm\, \partial_{y^k}\hat{\phi}(y)\Big) \\ =\ & 0\end{aligned}$$

This must equal zero to satisfy the Poincaré algebra. Note that for general $\hat{\phi}(x)$ and $\hat{\pi}(x)$, there is no way for the $\big[\hat{\phi}(y), \hat{\phi}(x)\big]$ and $\big[\hat{\pi}(y), \hat{\pi}(x)\big]$ terms to cancel in general. Therefore, it must be that

$$\big[\hat{\phi}(y), \hat{\phi}(x)\big]_\pm = \big[\hat{\pi}(y), \hat{\pi}(x)\big]_\pm = 0 \tag{96}$$

The $\big[\hat{\pi}(y), \hat{\phi}(x)\big]$ terms, on the other hand, will cancel if we hypothesize that

$$\big[\hat{\pi}(y), \hat{\phi}(x)\big]_\pm = C_\pm\, \delta^3(\vec{x} - \vec{y})\, \hat{\mathbb{1}} \tag{97}$$

where $C_\pm$ is some complex constant. This is easily demonstrated by integrating out the δ-functions, and integrating by parts again.

$$\begin{aligned}&[\hat{P}_j, \hat{P}_k] \\ =\ & \int d^3x\, d^3y \Big(\pm \big(\partial_{y^k}\hat{\pi}(y)\big)\, \big[\hat{\pi}(y), \hat{\phi}(x)\big]_\pm\, \partial_{x^j}\hat{\phi}(x)\ \mp\ \big(\partial_{x^j}\hat{\pi}(x)\big)\, \big[\hat{\pi}(y), \hat{\phi}(x)\big]_\pm\, \partial_{y^k}\hat{\phi}(y)\Big) \\ =\ & C_\pm \int d^3x\, d^3y\, \delta^3(\vec{x} - \vec{y})\, \Big(\pm \big(\partial_{y^k}\hat{\pi}(y)\big)\, \partial_{x^j}\hat{\phi}(x)\ \mp\ \big(\partial_{x^j}\hat{\pi}(x)\big)\, \partial_{y^k}\hat{\phi}(y)\Big) \\ =\ & C_\pm \int d^3x \Big(\pm \big(\partial_k\hat{\pi}(x)\big)\, \partial_j\hat{\phi}(x)\ \mp\ \big(\partial_j\hat{\pi}(x)\big)\, \partial_k\hat{\phi}(x)\Big) \\ =\ & C_\pm \int d^3x \Big(\mp \cancel{\big(\partial_j\partial_k\hat{\pi}(x)\big)\, \hat{\phi}(x)}\ \pm\ \cancel{\big(\partial_k\partial_j\hat{\pi}(x)\big)\, \hat{\phi}(x)}\Big) = 0\end{aligned}$$

We have derived the equal time commutation relations for the real scalar quantum field, except that we do not know the constant $C_\pm$ and it is ambiguous whether the commutation relations should involve commutators or anticommutators. We can discover the value of the constant $C_\pm$ by examining another part of the Poincaré algebra. Equation 93 from the Poincaré algebra says

$$\big[\hat{M}_{jk}, \hat{P}_\ell\big] = -i\, \eta_{j\ell}\, \hat{P}_k + i\, \eta_{k\ell}\, \hat{P}_j$$

$\hat{M}_{jk}$ is given by equations 75 and 78.

$$\hat{M}_{jk}(x^0) = \int d^3x \Big(x_j\, \hat{\pi}(x)\, \partial_k\hat{\phi}(x) - x_k\, \hat{\pi}(x)\, \partial_j\hat{\phi}(x)\Big)$$

We will calculate the following equal time commutator ($x^0 = y^0$) in terms of $\hat{\phi}$ and $\hat{\pi}$.

$$\begin{aligned}
&\big[\hat{M}_{jk}(x^0), \hat{P}_\ell(y^0)\big] \\
= &\int d^3x\, d^3y \big[x_j\, \hat{\pi}(x)\, \partial_{x^k}\hat{\phi}(x) - x_k\, \hat{\pi}(x)\, \partial_{x^j}\hat{\phi}(x), \hat{\pi}(y)\, \partial_{y^\ell}\hat{\phi}(y)\big] \\
= &\int d^3x\, d^3y \big[x_j\, \hat{\pi}(x)\, \partial_{x^k}\hat{\phi}(x), \hat{\pi}(y)\, \partial_{y^\ell}\hat{\phi}(y)\big] - \big(j \leftrightarrow k\big)
\end{aligned}$$

Consider the first term. Note that x_j is not an operator and can therefore be commuted past everything.

$$\begin{aligned}
&\int d^3x\, d^3y \big[x_j\, \hat{\pi}(x)\, \partial_{x^k}\hat{\phi}(x), \hat{\pi}(y)\, \partial_{y^\ell}\hat{\phi}(y)\big] \\
= &\int d^3x\, d^3y\, x_j \Big(\hat{\pi}(x)\, \partial_{x^k}\hat{\phi}(x)\, \hat{\pi}(y)\, \partial_{y^\ell}\hat{\phi}(y) - \hat{\pi}(y)\, \partial_{y^\ell}\, \underbrace{\hat{\phi}(y)\, \hat{\pi}(x)}_{\text{commute}}\, \partial_{x^k}\hat{\phi}(x)\Big) \\
= &\int d^3x\, d^3y\, x_j \Big(\hat{\pi}(x)\, \partial_{x^k}\hat{\phi}(x)\, \hat{\pi}(y)\, \partial_{y^\ell}\hat{\phi}(y) - \hat{\pi}(y)\, \partial_{y^\ell}\big[\hat{\phi}(y), \hat{\pi}(x)\big]_\pm\, \partial_{x^k}\hat{\phi}(x) \\
&\pm \underbrace{\hat{\pi}(y)\, \hat{\pi}(x)}_{\text{commute}}\, \underbrace{\partial_{y^\ell}\hat{\phi}(y)\, \partial_{x^k}\hat{\phi}(x)}_{\text{commute}}\Big)
\end{aligned}$$

Note that we have already concluded that $\big[\hat{\phi}(y), \hat{\phi}(x)\big]_\pm = \big[\hat{\pi}(y), \hat{\pi}(x)\big]_\pm = 0$, from equation 96.

$$\begin{aligned}
= &\int d^3x\, d^3y\, x_j \Big(\hat{\pi}(x)\, \partial_{x^k}\hat{\phi}(x)\, \hat{\pi}(y)\, \partial_{y^\ell}\hat{\phi}(y) - \hat{\pi}(y)\, \partial_{y^\ell}\big[\hat{\phi}(y), \hat{\pi}(x)\big]_\pm\, \partial_{x^k}\hat{\phi}(x) \\
&\pm \hat{\pi}(x)\, \underbrace{\hat{\pi}(y)\, \partial_{x^k}\hat{\phi}(x)}_{\text{commute}}\, \partial_{y^\ell}\hat{\phi}(y)\Big) \\
= &\int d^3x\, d^3y\, x_j \Big(\cancel{\hat{\pi}(x)\, \partial_{x^k}\hat{\phi}(x)\, \hat{\pi}(y)\, \partial_{y^\ell}\hat{\phi}(y)} - \hat{\pi}(y)\, \partial_{y^\ell}\big[\hat{\phi}(y), \hat{\pi}(x)\big]_\pm\, \partial_{x^k}\hat{\phi}(x) \\
&\pm \hat{\pi}(x)\, \partial_{x^k}\big[\hat{\pi}(y), \hat{\phi}(x)\big]_\pm\, \partial_{y^\ell}\hat{\phi}(y) - \cancel{\hat{\pi}(x)\, \partial_{x^k}\hat{\phi}(x)\, \hat{\pi}(y)\, \partial_{y^\ell}\hat{\phi}(y)}\Big)
\end{aligned}$$

Flip the first commutator and integrate the remaining terms by parts.

$$\begin{aligned}
= &\int d^3x\, d^3y\, x_j \Big(\mp \hat{\pi}(y)\, \partial_{y^\ell}\big[\hat{\pi}(x), \hat{\phi}(y)\big]_\pm\, \partial_{x^k}\hat{\phi}(x)\, \pm\, \hat{\pi}(x)\, \partial_{x^k}\big[\hat{\pi}(y), \hat{\phi}(x)\big]_\pm\, \partial_{y^\ell}\hat{\phi}(y)\Big) \\
= &\int d^3x\, d^3y \Big(\pm x_j\, \big(\partial_{y^\ell}\hat{\pi}(y)\big)\, \big[\hat{\pi}(x), \hat{\phi}(y)\big]_\pm\, \partial_{x^k}\hat{\phi}(x)\, \mp\, \partial_{x^k}\big(x_j\, \hat{\pi}(x)\big)\, \big[\hat{\pi}(y), \hat{\phi}(x)\big]_\pm\, \partial_{y^\ell}\hat{\phi}(y)\Big) \\
= &\; C_\pm \int d^3x\, d^3y\, \delta^3(\vec{x} - \vec{y}) \Big(\pm x_j\, \big(\partial_{y^\ell}\hat{\pi}(y)\big)\, \partial_{x^k}\hat{\phi}(x)\, \mp\, \partial_{x^k}\big(x_j\, \hat{\pi}(x)\big)\, \partial_{y^\ell}\hat{\phi}(y)\Big) \\
= &\; C_\pm \int d^3x \Big(\pm x_j\, \big(\partial_\ell\hat{\pi}(x)\big)\, \partial_k\hat{\phi}(x)\, \mp\, \partial_k\big(x_j\, \hat{\pi}(x)\big)\, \partial_\ell\hat{\phi}(x)\Big)
\end{aligned}$$

$$\partial_k\big(x_j\,\hat{\pi}(x)\big) = \underbrace{\frac{\partial x_j}{\partial x^k}}_{\eta_{jk}}\,\hat{\pi}(x) + x_j\,\partial_k\hat{\pi}(x)$$

$$= \quad C_\pm \int d^3x \Big(\pm\, x_j\,\big(\partial_\ell \hat{\pi}(x)\big)\,\partial_k\hat{\phi}(x) \mp \eta_{jk}\,\hat{\pi}(x)\,\partial_\ell\hat{\phi}(x) \mp x_j\,\partial_k\hat{\pi}(x)\partial_\ell\hat{\phi}(x)\Big)$$

Now, integrate the first and third terms by parts.

$$\begin{aligned} &= \quad C_\pm \int d^3x \Big(\mp\,\hat{\pi}(x)\,\partial_\ell\big(x_j\,\partial_k\hat{\phi}(x)\big) \mp \eta_{jk}\,\hat{\pi}(x)\,\partial_\ell\hat{\phi}(x) \pm \hat{\pi}(x)\,\partial_k\big(x_j\partial_\ell\hat{\phi}(x)\big)\Big) \\ &= \quad C_\pm \int d^3x \Big(\mp\,\hat{\pi}(x)\,\eta_{j\ell}\,\partial_k\hat{\phi}(x) \mp \cancel{x_j\,\hat{\pi}(x)\,\partial_\ell\partial_k\hat{\phi}(x)} \\ &\qquad \mp \cancel{\eta_{jk}\,\hat{\pi}(x)\,\partial_\ell\hat{\phi}(x)} \pm \cancel{\hat{\pi}(x)\,\eta_{jk}\,\partial_\ell\hat{\phi}(x)} \pm \cancel{x_j\,\hat{\pi}(x)\partial_k\partial_\ell\hat{\phi}(x)}\Big) \end{aligned}$$

We see that we are left with the definition of $\hat{P}_k$.

$$\begin{aligned} \int d^3x\,d^3y\big[x_j\,\hat{\pi}(x)\,\partial_{x^k}\hat{\phi}(x), \hat{\pi}(y)\,\partial_{y^\ell}\hat{\phi}(y)\big] \quad &= \quad \mp\, C_\pm\,\eta_{j\ell} \underbrace{\int d^3x\,\hat{\pi}(x)\,\partial_k\hat{\phi}(x)}_{\hat{P}_k} \\ &= \quad \mp\, C_\pm\,\eta_{j\ell}\,\hat{P}_k \end{aligned}$$

Similarly for the $(j \leftrightarrow k)$ term:

$$\int d^3x\,d^3y\big[x_k\,\hat{\pi}(x)\,\partial_{x^j}\hat{\phi}(x), \hat{\pi}(y)\,\partial_{y^\ell}\hat{\phi}(y)\big] = \mp\, C_\pm\,\eta_{k\ell}\,\hat{P}_j$$

Putting these two terms together gives $\big[\hat{M}_{jk}, \hat{P}_\ell\big]$.

$$\big[\hat{M}_{jk}, \hat{P}_\ell\big] = \mp\, C_\pm\big(\eta_{j\ell}\,\hat{P}_k - \eta_{k\ell}\,\hat{P}_j\big)$$

This agrees with equation 93 from the Poincaré algebra only if

$$C_\pm = \pm i \tag{98}$$

We are still left wondering whether we should use commutators or anticommutators. At this point, there is no way to resolve this ambiguity. For the sake of further discussion, we will correctly choose to use commutators for the real scalar field. We will revisit how this ambiguity is resolved when we discuss the spin-statistics theorem in section 3.3.7.

Summarizing our conclusions, the **equal time commuation relations** for the real scalar field are

$$\boxed{\begin{aligned}\left[\hat\phi(x),\hat\phi(y)\right] &= \left[\hat\pi(x),\hat\pi(y)\right] = 0 && (99)\\ \left[\hat\pi(x),\hat\phi(y)\right] &= -i\,\delta^3(\vec{x}-\vec{y})\,\hat{\mathbb{1}} && (100)\end{aligned}}$$

Let us now investigate some of the consequences of these equal time commutation relations. Plugging $\hat\phi(x)$ into the Heisenberg equation of motion, equation 91 gives

$$\begin{aligned}\partial_0\hat\phi(x) &= i\big[\hat H,\hat\phi(x)\big]\\ &= \frac{i}{2}\Big[\int d^3y\Big(\hat\pi^2(y)+\big(\nabla_y\hat\phi(y)\big)^2+m^2\hat\phi^2(y)\Big),\hat\phi(x)\Big]\\ &= \frac{i}{2}\int d^3y\big[\hat\pi^2(y),\hat\phi(x)\big]\\ &= \frac{i}{2}\int d^3y\Big(\hat\pi(y)\,\hat\pi(y)\,\hat\phi(x)-\hat\phi(x)\,\hat\pi(y)\,\hat\pi(y)\Big)\\ &= \frac{i}{2}\int d^3y\Big(\cancel{\hat\pi(y)\,\hat\phi(x)\,\hat\pi(y)}+\hat\pi(y)\,\big[\hat\pi(y),\hat\phi(x)\big]\\ &\quad-\cancel{\hat\pi(y)\,\hat\phi(x)\,\hat\pi(y)}+\big[\hat\pi(y),\hat\phi(x)\big]\,\hat\pi(y)\Big)\\ &= i\int d^3y\,\hat\pi(y)\,(-i)\,\delta^3(\vec{x}-\vec{y})=\hat\pi(x)\end{aligned}$$

$$\therefore\qquad \boxed{\hat\pi(x)=\dot{\hat\phi}(x)} \qquad (101)$$

Therefore, the conjugate momentum has the same expression we derived in equation 73 for the classical case. Now plug $\hat\pi(x)$ into the Heisenberg equation of motion.

$$\partial_0\hat\pi(x)=\frac{i}{2}\Big[\int d^3y\Big(\cancelto{0}{\hat\pi^2(y)}+\big(\nabla_y\hat\phi(y)\big)^2+m^2\hat\phi^2(y)\Big),\hat\pi(x)\Big]$$

Integrate the second term by parts once.

$$\begin{aligned}\partial_0\hat\pi(x) &= \frac{i}{2}\int d^3y\big[\hat\phi(y)\,(-\nabla_y^2+m^2)\,\hat\phi(y),\hat\pi(x)\big]\\ &= \frac{i}{2}\int d^3y\Big(\hat\phi(y)\,(-\nabla_y^2+m^2)\,\hat\phi(y)\,\hat\pi(x)-\hat\pi(x)\,\hat\phi(y)\,(-\nabla_y^2+m^2)\,\hat\phi(y)\Big)\\ &= \frac{i}{2}\int d^3y\Big(\cancel{\hat\phi(y)\,(-\nabla_y^2+m^2)\,\hat\pi(x)\,\hat\phi(y)}-\hat\phi(y)\,(-\nabla_y^2+m^2)\,\big[\hat\pi(x),\hat\phi(y)\big]\\ &\quad-\cancel{\hat\phi(y)\,\hat\pi(x)\,(-\nabla_y^2+m^2)\,\hat\phi(y)}-\big[\hat\pi(x),\hat\phi(y)\big]\,(-\nabla_y^2+m^2)\,\hat\phi(y)\Big)\end{aligned}$$

Integrate the second term by parts twice.

$$\begin{aligned}\partial_0\hat\pi(x) &= -i\int d^3y\big[\hat\pi(x),\hat\phi(y)\big]\,(-\nabla_y^2+m^2)\,\hat\phi(y)\\ &= -\int d^3y\,\delta^3(\vec{x}-\vec{y})\,(-\nabla_y^2+m^2)\,\hat\phi(y)\\ \ddot{\hat\phi}(x) &= -(-\nabla^2+m^2)\,\hat\phi(x)\end{aligned}$$

$$\Rightarrow \qquad \ddot{\hat{\phi}}(x) - \nabla^2 \hat{\phi}(x) + m^2 \hat{\phi}(x) = 0$$
$$\therefore \qquad \boxed{\left(\partial^2 + m^2\right) \hat{\phi} = 0} \tag{102}$$

Which is the Klein-Gordon equation, the same equation of motion we derived for the classical real scalar field. We have just seen an example of the general fact that *quantum operators satisfy the classical equations of motion.*

3.3.2 Creation and Annihilation Operators

Similar to equation 85, we expand $\hat{\phi}(x)$ in terms of plane-waves, where the "coefficients" in the expansion are now themselves operators.

$$\boxed{\hat{\phi}(x) = \int \frac{d^3k}{(2\pi)^3} \left(\hat{a}(k)\, f_k(x) + \hat{a}^\dagger(k)\, f_k^*(x)\right)} \tag{103}$$

Similar to equation 85, we can project out $\hat{a}(k)$.

$$\boxed{\hat{a}(k) = i \int d^3x\, e^{i\,k\cdot x}\, \overleftrightarrow{\partial_0}\, \hat{\phi}(x)} \tag{104}$$

$$\boxed{\hat{a}(k) = \int d^3x\, e^{i\,k\cdot x} \left(\omega_k\, \hat{\phi}(x) + i\, \hat{\pi}(x)\right)} \tag{105}$$

We would now like to derive the algebra for the $\hat{a}(k)$ operators. Consider

$$\begin{aligned}
\left[\hat{a}(k), \hat{a}(k')\right] &= \int d^3x\, d^3y\, f_k^*(x)\, f_{k'}^*(y)\, \left[\omega_k\, \hat{\phi}(x) + i\, \hat{\pi}(x), \omega_{k'}\, \hat{\phi}(y) + i\, \hat{\pi}(y)\right] \\
&= \int d^3x\, d^3y\, f_k^*(x)\, f_{k'}^*(y) \Big(\omega_k\, \omega_{k'}\, \cancelto{0}{\left[\hat{\phi}(x), \hat{\phi}(y)\right]} - \cancelto{0}{\left[\hat{\pi}(x), \hat{\pi}(y)\right]} \\
&\quad + i\, \omega_k\, \underbrace{\left[\hat{\phi}(x), \hat{\pi}(y)\right]}_{i\,\delta^3(\vec{x}-\vec{y})\,\hat{\mathbb{1}}} + i\, \omega_{k'}\, \underbrace{\left[\hat{\pi}(x), \hat{\phi}(y)\right]}_{-i\,\delta^3(\vec{x}-\vec{y})\,\hat{\mathbb{1}}}\Big) \\
&= \int d^3x\, d^3y\, f_k^*(x)\, f_{k'}^*(x)\, (\omega_{k'} - \omega_k)\, \hat{\mathbb{1}} \\
&= \frac{\omega_{k'} - \omega_k}{2\sqrt{\omega_k\, \omega_{k'}}} \int d^3x\, e^{i(k+k')\cdot x}\, \hat{\mathbb{1}} \\
&= \frac{\omega_{k'} - \omega_k}{2\sqrt{\omega_k\, \omega_{k'}}}\, e^{i(\omega_k + \omega_{k'})x^0} \underbrace{\int d^3x\, e^{-i(\vec{k}+\vec{k}')\cdot x}}_{(2\pi)^3\,\delta^3\left(\vec{k}+\vec{k}'\right)}\, \hat{\mathbb{1}} = 0
\end{aligned}$$

Because $\vec{k} = -\vec{k'} \;\Rightarrow\; \omega_k = \omega_{k'}$.

$$\Rightarrow \qquad \boxed{\left[\hat{a}(k), \hat{a}(k')\right] = \left[\hat{a}^\dagger(k), \hat{a}^\dagger(k')\right] = 0} \tag{106}$$

Now consider

$$\begin{aligned}
\left[\hat{a}(k), \hat{a}^\dagger(k')\right] &= \int d^3x\, d^3y\, f_k^*(x)\, f_{k'}(y)\, \left[\omega_k\, \hat{\phi}(x) + i\, \hat{\pi}(x), \omega_{k'}\, \hat{\phi}(y) - i\, \hat{\pi}(y)\right] \\
&= \int d^3x\, d^3y\, f_k^*(x)\, f_{k'}(y) \Big(\omega_k\, \omega_{k'}\, \cancelto{0}{\left[\hat{\phi}(x), \hat{\phi}(y)\right]} + \cancelto{0}{\left[\hat{\pi}(x), \hat{\pi}(y)\right]} \\
&\quad - i\, \omega_k\, \underbrace{\left[\hat{\phi}(x), \hat{\pi}(y)\right]}_{i\,\delta^3(\vec{x}-\vec{y})\hat{\mathbb{1}}} + i\, \omega_{k'}\, \underbrace{\left[\hat{\pi}(x), \hat{\phi}(y)\right]}_{-i\,\delta^3(\vec{x}-\vec{y})\,\hat{\mathbb{1}}} \Big) \\
&= \int d^3x\, d^3y\, f_k^*(x)\, f_{k'}(x)\, (\omega_{k'} + \omega_k)\, \hat{\mathbb{1}} \\
&= \frac{\omega_{k'} + \omega_k}{2\sqrt{\omega_k\, \omega_{k'}}}\, e^{i(\omega_k - \omega_{k'})x^0} \underbrace{\int d^3x\, e^{-i(\vec{k}-\vec{k}')\cdot x}}_{(2\pi)^3\,\delta^3\left(\vec{k}-\vec{k}'\right)}\, \hat{\mathbb{1}} \\
&= \frac{2\,\omega_k}{2\,\omega_k}\, e^0\, (2\pi)^3\, \delta^3\left(\vec{k} - \vec{k}'\right)\, \hat{\mathbb{1}}
\end{aligned}$$

$$\therefore \quad \boxed{\left[\hat{a}(k), \hat{a}^\dagger(k')\right] = (2\pi)^3\, \delta^3\left(\vec{k} - \vec{k}'\right)\, \hat{\mathbb{1}}} \tag{107}$$

We will come to see why $\hat{a}^\dagger(k)$ and $\hat{a}(k)$ are known as **creation and annihilation operators** respectively. But first, we define the **number operators**, $\hat{N}(k)$ as

$$\hat{N}(k) \equiv \hat{a}^\dagger(k)\, \hat{a}(k) \tag{108}$$

Note that $\hat{N}(k)$ are hermitian.

$$\hat{N}^\dagger(k) = \hat{N}(k)$$

For the sake of clarity in the following discussion, let us discretize the four-momentum k, to have enumerable values k_j. Our conclusions will be true in the continuous k limit. In this case, the creation and annihilation operators have the following algebra.

$$\left[\hat{a}(k_i), \hat{a}^\dagger(k_j)\right] = \delta_{ij}\, \hat{\mathbb{1}}$$

Let $|n_j\rangle$ denote the eigenstate of the $\hat{N}(k_j)$ operator, with eigenvalue n_j.

$$\hat{N}(k_j)\, |n_j\rangle = n_j\, |n_j\rangle$$

Let

$$|\tilde{n}_j\rangle \equiv \hat{a}(k_j)\, |n_j\rangle$$

Then

$$\begin{aligned}
\hat{N}(k_j)\,|\tilde{n}_j\rangle &= \hat{a}^\dagger(k_j)\,\hat{a}(k_j)\,\hat{a}(k_j)\,|n_j\rangle \\
&= \left(\hat{a}(k_j)\,\hat{a}^\dagger(k_j) - \left[\hat{a}(k_j), \hat{a}^\dagger(k_j)\right]\right)\hat{a}(k_j)\,|n_j\rangle \\
&= \left(\hat{a}(k_j)\,\hat{a}^\dagger(k_j) - \hat{\mathbb{1}}\right)\hat{a}(k_j)\,|n_j\rangle \\
&= \hat{a}(k_j)\left(\hat{a}(k_j)\,\hat{a}^\dagger(k_j) - \hat{\mathbb{1}}\right)|n_j\rangle \\
&= \hat{a}(k_j)\,(n_j - 1)\,|n_j\rangle \\
&= (n_j - 1)\,|\tilde{n}_j\rangle
\end{aligned}$$

Therefore, $\hat{a}(k_j)\,|n_j\rangle$ is an eigenstate of $\hat{N}(k_j)$, with eigenvalue one less than that of $|n_j\rangle$.

$$\Rightarrow \qquad \hat{a}(k_j)\,|n_j\rangle \propto |n_j - 1\rangle$$

We say that "$\hat{a}(k_j)$ *lowers* eigenstates". An analogous argument shows that $\hat{a}^\dagger(k_j)$ *raises* eigenstates.

$$\hat{a}^\dagger(k_j)\,|n_j\rangle \propto |n_j + 1\rangle$$

Assuming that $|n_j\rangle$ are normalized such that $\langle n_j|n_j\rangle = 1$, the following argument shows that the eigenvalues n_j must be nonnegative.

$$n_j = \langle n_j|\hat{N}(k_j)|n_j\rangle = \langle n_j|\hat{a}^\dagger(k_j)\,\hat{a}(k_j)|n_j\rangle = \langle \tilde{n}_j|\tilde{n}_j\rangle \geq 0$$

We hypothesize that there is a lowest eigenstate, $|0\rangle$, with eigenvalue zero[5].

$$\hat{a}(k_j)\,|0\rangle = 0 \qquad \forall\, k_j \tag{109}$$

We call this state the "**vacuum**."

We will now find the normalization constant needed to keep raised and lowered states normalized. Let

$$\hat{a}(k_j)\,|n_j\rangle = C\;|n_j - 1\rangle$$

Then

$$n_j = \langle n_j|\hat{N}(k_j)|n_j\rangle = \langle n_j|\hat{a}^\dagger(k_j)\,\hat{a}(k_j)|n_j\rangle = C^2\;\langle n_j - 1|n_j - 1\rangle = C^2$$

$$\Rightarrow \qquad C = \sqrt{n_j}$$

$$\therefore \qquad \hat{a}(k_j)\,|n_j\rangle = \sqrt{n_j}\;|n_j - 1\rangle \tag{110}$$

[5] Note that $|0\rangle$ is *not* the zero vector. $|\psi\rangle + |0\rangle \neq |\psi\rangle$

Similarly, one can show that

$$\hat{a}^\dagger(k_j)\,|n_j\rangle = \sqrt{n_j+1}\;|n_j+1\rangle \tag{111}$$

Using this, we can build up all the states from the vacuum by raising.

$$|n_j\rangle \propto \left(\hat{a}^\dagger(k_j)\right)^n |0\rangle$$

$$\left(\hat{a}^\dagger(k_j)\right)^n |0\rangle = \left(\hat{a}^\dagger(k_j)\right)^{n-1}\sqrt{1}\;|1_j\rangle = \left(\hat{a}^\dagger(k_j)\right)^{n-2}\sqrt{2}\;|2_j\rangle = \ldots = \sqrt{n!}\;|n_j\rangle$$

$$\Rightarrow \qquad |n_j\rangle = \frac{1}{\sqrt{n!}}\left(\hat{a}^\dagger(k_j)\right)^n |0\rangle \tag{112}$$

Since each raising increases the eigenvalue n_j by one, and the vacuum has eigenvalue zero, we know that *the eigenvalues of the number operators are all nonnegative integers.* The fact that $\left[\hat{a}(k_i), \hat{a}(k_j)\right] = 0$ implies that the $\hat{N}(k_j)$ commute as well.

$$\left[\hat{N}(k_i), \hat{N}(k_j)\right] = 0$$

Which implies that $\hat{N}(k_j)$ have simultaneous eigenstates. We can label the states by the eigenvalues of each $\hat{N}(k_j)$.

$$\begin{aligned} |n_1, n_2, \ldots\rangle &= \frac{1}{\sqrt{n_1!}}\left(\hat{a}^\dagger(k_1)\right)^{n_1}\frac{1}{\sqrt{n_2!}}\left(\hat{a}^\dagger(k_2)\right)^{n_2}\ldots|0\rangle \\ &= \prod_j \frac{1}{\sqrt{n_j!}}\left(\hat{a}^\dagger(k_j)\right)^{n_j}|0\rangle \end{aligned}$$

In the limit of continuous k, we can label a state by a single integer-valued function, $n(k)$.

$$\boxed{|n(k)\rangle = \prod_k \frac{1}{\sqrt{n(k)!}}\left(\hat{a}^\dagger(k)\right)^{n(k)}|0\rangle} \tag{113}$$

So far, we have not discussed the interpretation of these states $|n(k)\rangle$. We will come to see that $|n(k)\rangle$ is a state with $n(k)$ particles with momentum k, and the number operators, $\hat{N}(k)$, count the number of particles with momentum k. In order to motivate this interpretation, we should first investigate the total energy and momentum of these states.

3.3.3 Energy Eigenstates

Recall that the Klein-Gordon Hamiltonian (equation 74) is

$$\hat{H} = \frac{1}{2}\int d^3x\left(\hat{\pi}^2(x) + (\vec{\nabla}\hat{\phi}(x))^2 + m^2\,\hat{\phi}^2(x)\right)$$

Also, recall our plane-wave expansion for $\hat{\phi}(x)$ (equation 103).

$$\hat{\phi}(x) = \int \frac{d^3k}{(2\pi)^3}\Big(\hat{a}(k)\, f_k(x) + \hat{a}^\dagger(k)\, f_k^*(x)\Big)$$

We want to express $\hat{H}$ in terms of $\hat{\phi}(x)$ and $\hat{\pi}(x)$, but first we need to calculate $\hat{\pi}^2(x)$, $(\vec{\nabla}\hat{\phi}(x))^2$, and $\hat{\phi}^2(x)$.

$$\hat{\pi}(x) = \dot{\hat{\phi}}(x) = i\int \frac{d^3k}{(2\pi)^3}\, \omega_k\, \Big(-\hat{a}(k)\, f_k(x) + \hat{a}^\dagger(k)\, f_k^*(x)\Big)$$

$$\vec{\nabla}\hat{\phi}(x) = i\int \frac{d^3k}{(2\pi)^3}\, \vec{k}\, \Big(\hat{a}(k)\, f_k(x) - \hat{a}^\dagger(k)\, f_k^*(x)\Big)$$

$$\begin{aligned}
\hat{\pi}^2(x) &= -\int \frac{d^3k}{(2\pi)^3}\,\frac{d^3k'}{(2\pi)^3}\,\omega_k\,\omega_{k'}\,\Big(-\hat{a}(k)\, f_k(x) + \hat{a}^\dagger(k)\, f_k^*(x)\Big)\Big(-\hat{a}(k')\, f_{k'}(x) + \hat{a}^\dagger(k')\, f_{k'}^*(x)\Big)\\
&= -\int \frac{d^3k}{(2\pi)^3}\,\frac{d^3k'}{(2\pi)^3}\,\omega_k\,\omega_{k'}\,\Big(\hat{a}(k)\,\hat{a}(k')\, f_k(x)\, f_{k'}(x) + \hat{a}^\dagger(k)\,\hat{a}^\dagger(k')\, f_k^*(x)\, f_{k'}^*(x)\\
&\quad -\hat{a}^\dagger(k)\,\hat{a}(k')\, f_k^*(x)\, f_{k'}(x) - \hat{a}(k)\,\hat{a}^\dagger(k')\, f_k(x)\, f_{k'}^*(x)\Big)
\end{aligned}$$

Similarly

$$\begin{aligned}
|\vec{\nabla}\hat{\phi}(x)|^2 &= -\int \frac{d^3k}{(2\pi)^3}\,\frac{d^3k'}{(2\pi)^3}\,\vec{k}\cdot\vec{k'}\,\Big(\hat{a}(k)\,\hat{a}(k')\, f_k(x)\, f_{k'}(x) + \hat{a}^\dagger(k)\,\hat{a}^\dagger(k')\, f_k^*(x)\, f_{k'}^*(x)\\
&\quad -\hat{a}^\dagger(k)\,\hat{a}(k')\, f_k^*(x)\, f_{k'}(x) - \hat{a}(k)\,\hat{a}^\dagger(k')\, f_k(x)\, f_{k'}^*(x)\Big)
\end{aligned}$$

and[6]

$$\begin{aligned}
m^2\,\hat{\phi}^2(x) &= \int \frac{d^3k}{(2\pi)^3}\,\frac{d^3k'}{(2\pi)^3}\,\vec{k}\cdot\vec{k'}\,\Big(\hat{a}(k)\,\hat{a}(k')\, f_k(x)\, f_{k'}(x) + \hat{a}^\dagger(k)\,\hat{a}^\dagger(k')\, f_k^*(x)\, f_{k'}^*(x)\\
&\quad +\hat{a}^\dagger(k)\,\hat{a}(k')\, f_k^*(x)\, f_{k'}(x) + \hat{a}(k)\,\hat{a}^\dagger(k')\, f_k(x)\, f_{k'}^*(x)\Big)
\end{aligned}$$

Therefore, putting these together gives

$$\begin{aligned}
\hat{H} &= \frac{1}{2}\int \frac{d^3k}{(2\pi)^3}\,\frac{d^3k'}{(2\pi)^3}\,d^3x\Big[\\
&\Big(-\omega_k\,\omega_{k'} - \vec{k}\cdot\vec{k'} + m^2\Big)\Big(\hat{a}(k)\,\hat{a}(k')\, f_k(x)\, f_{k'}(x) + \hat{a}^\dagger(k)\,\hat{a}^\dagger(k')\, f_k^*(x)\, f_{k'}^*(x)\Big)\\
&+\Big(+\omega_k\,\omega_{k'} + \vec{k}\cdot\vec{k'} + m^2\Big)\Big(\hat{a}^\dagger(k)\,\hat{a}(k')\, f_k^*(x)\, f_{k'}(x) + \hat{a}(k)\,\hat{a}^\dagger(k')\, f_k(x)\, f_{k'}^*(x)\Big)\Big]
\end{aligned}$$

[6] Note the signs of the terms of $\hat{\pi}^2(x)$ and $|\vec{\nabla}\hat{\phi}(x)|^2$ are the same, but $m^2\,\hat{\phi}^2(x)$ has all positive terms.

We then do the following integrals

$$\int d^3x \; f_k(x)\; f^*_{k'}(x) \;\; = \;\; \frac{e^{-i(\omega_k - \omega_{k'})x^0}}{2\,\sqrt{\omega_k\,\omega_{k'}}} \int d^3x \; e^{i(\vec{k}-\vec{k'})\cdot\vec{x}}$$
$$= \;\; \frac{1}{2\,\omega_k}\,(2\pi)^3\,\delta^3\left(\vec{k}-\vec{k'}\right)$$

$$\int d^3x \; f_k(x)\; f_{k'}(x) \;\; = \;\; \frac{e^{-i(\omega_k + \omega_{k'})x^0}}{2\,\sqrt{\omega_k\,\omega_{k'}}} \int d^3x \; e^{i(\vec{k}+\vec{k'})\cdot\vec{x}}$$
$$= \;\; \frac{e^{-i\,2\,\omega_k\,x^0}}{2\,\omega_k}\,(2\pi)^3\,\delta^3\left(\vec{k}+\vec{k'}\right)$$

This leaves

$$\begin{aligned}
\hat{H} \;\; &= \;\; \frac{1}{2}\int \frac{d^3k}{(2\pi)^3}\,\frac{d^3k'}{\cancel{(2\pi)^3}}\Bigg[\\
&\left(-\,\omega_k\,\omega_{k'} - \vec{k}\cdot\vec{k'} + m^2\right)\left(\hat{a}(k)\,\hat{a}(k')\,e^{-i\,2\,\omega_k\,x^0} + \hat{a}^\dagger(k)\,\hat{a}^\dagger(k')\,e^{+i\,2\,\omega_k\,x^0}\right)\frac{\cancel{(2\pi)^3}\,\delta^3\left(\vec{k}+\vec{k'}\right)}{2\,\omega_k} \\
&+\left(+\,\omega_k\,\omega_{k'} + \vec{k}\cdot\vec{k'} + m^2\right)\left(\hat{a}^\dagger(k)\,\hat{a}(k') + \hat{a}(k)\,\hat{a}^\dagger(k')\right)\frac{\cancel{(2\pi)^3}\,\delta^3\left(\vec{k}-\vec{k'}\right)}{2\,\omega_k}\Bigg] \\
&= \;\; \frac{1}{2}\int \frac{d^3k}{(2\pi)^3}\Bigg[\\
&\underbrace{\left(-\,\omega_k^2 + \vec{k}^2 + m^2\right)}_{0}\left(\hat{a}(k)\,\hat{a}(-k)\,e^{-i\,2\,\omega_k\,x^0} + \hat{a}^\dagger(k)\,\hat{a}^\dagger(-k)\,e^{+i\,2\,\omega_k\,x^0}\right)\frac{1}{2\,\omega_k} \\
&+\underbrace{\left(+\,\omega_k^2 + \vec{k}^2 + m^2\right)}_{2\,\omega_k^2}\left(\hat{a}^\dagger(k)\,\hat{a}(k) + \hat{a}(k)\,\hat{a}^\dagger(k)\right)\frac{1}{2\,\omega_k}\Bigg]
\end{aligned}$$

Therefore

$$\hat{H} = \frac{1}{2}\int \frac{d^3k}{(2\pi)^3}\,\omega_k\left(\hat{a}^\dagger(k)\,\hat{a}(k) + \hat{a}(k)\,\hat{a}^\dagger(k)\right)$$

Note that we have not yet used a commutation relation to commute operators. We have left them in order. We will now use our commutation relation for the creation and annihilation operators (equation 107) that we derived as a consequence of the equal time commutation relation for $\hat{\phi}$ and $\hat{\pi}$ (equation 100).

$$\begin{aligned}
\hat{H} &= \int \frac{d^3k}{(2\pi)^3}\,\omega_k \left(\hat{a}^\dagger(k)\,\hat{a}(k) + \frac{1}{2}\left[\hat{a}(k),\hat{a}^\dagger(k)\right]\right) \\
&= \int \frac{d^3k}{(2\pi)^3}\,\omega_k \left(\hat{a}^\dagger(k)\,\hat{a}(k) + \frac{1}{2}\,\delta^3(\vec{k}-\vec{k})\,\hat{\mathbb{1}}\right) \\
&= \int \frac{d^3k}{(2\pi)^3}\,\omega_k \left(\hat{N}(k) + \frac{1}{2}\,\delta_k^3(0)\,\hat{\mathbb{1}}\right) \qquad (114)
\end{aligned}$$

The Hamiltonian commutes with the number operators, because the number operators commute with themselves. Therefore, the Hamiltonian and the number operators have simultaneous eigenstates. Similarly, using equations 75, 99, 100, 103, 106, and 107, one can show a similar expression for the three-momentum.

$$\hat{\vec{P}} = \int \frac{d^3k}{(2\pi)^3}\,\vec{k}\left(\hat{N}(k) + \frac{1}{2}\,\delta_k^3(0)\,\hat{\mathbb{1}}\right) \qquad (115)$$

Therefore

$$\hat{P}_\mu = \int \frac{d^3k}{(2\pi)^3}\,k_\mu\left(\hat{N}(k) + \frac{1}{2}\,\delta_k^3(0)\,\hat{\mathbb{1}}\right) \qquad (116)$$

From equation 114, we see that for any given $\vec{k}$, the energy steps in integer units of ω_k because we know the number operators have nonnegative integer eigenvalues. The second term involves a δ-function[7] always evaluated at its singular point. This term is an infinite constant energy that is there even when all the number operators have eigenvalue zero, called the "zero point energy." Evidently, any dynamics of the system will take the system to states with energies of various amounts of ω_k always added on top of this uninteresting infinite zero point energy. We can shift our energy scale so that the vacuum has zero energy, but we have to find a way to properly subtract the infinity. In order to do this, we introduce the concept of normal ordering.

3.3.4 Normal Ordering

Let us break $\hat{P}_\mu$ into its finite part and the infinite zero point energy. Let

$$\hat{P'}_\mu \equiv \int \frac{d^3k}{(2\pi)^3}\,k_\mu\,\hat{N}(k)$$

and

$$C_\mu \equiv \frac{1}{2}\int \frac{d^3k}{(2\pi)^3}\,k_\mu\,\delta_k^3(0)$$

Then

$$\hat{P}_\mu = \hat{P'}_\mu + C_\mu\,\hat{\mathbb{1}}$$

[7] The subscript k on the δ-function is to remind us that the δ-function is in momentum space.

$$\langle 0|\hat{P}_\mu|0\rangle = \cancelto{0}{\langle 0|\hat{P'}_\mu|0\rangle} + C_\mu \langle 0|0\rangle = C_\mu$$

because

$$\langle 0|\hat{N}(k)|0\rangle = 0$$

$$\Rightarrow \qquad \hat{P'}_\mu = \hat{P}_\mu - \langle 0|\hat{P}_\mu|0\rangle$$

Therefore, $\hat{P'}_\mu$ is the total four-momentum without the infinite zero point energy. If we define $\hat{M'}_{\mu\nu}$ as in equation 78, only with $\hat{P'}_\mu$ instead of $\hat{P}_\mu$, then $\hat{P'}_\mu$ and $\hat{M'}_{\mu\nu}$ still satisfy the Poincaré algebra because $\hat{P'}_\mu$ is the same operator as $\hat{P}_\mu$ shifted by a constant. The Heisenberg equation of motion is also unchanged because

$$\partial_\mu \hat{\phi}(x) = i[\hat{P}_\mu, \hat{\phi}(x)] = i[\hat{P'}_\mu, \hat{\phi}(x)]$$

$$\partial_\mu \hat{\pi}(x) = i[\hat{P}_\mu, \hat{\pi}(x)] = i[\hat{P'}_\mu, \hat{\pi}(x)]$$

which implies that the equations of motion are *unchanged* by using $\hat{P'}_\mu$ and $\hat{M'}_{\mu\nu}$ as the generators of the Poincaré transformations, instead of $\hat{P}_\mu$ and $\hat{M}_{\mu\nu}$.

Now we should formalize what we mean by $\hat{P'}_\mu$, other than $\hat{P}_\mu$ minus infinity. We will do this with normal ordering. Consider a real scalar field, $\hat{\phi}(x)$, with the plane-wave expansion given by equation 103. Let

$$\hat{\phi}(x) = \hat{\phi}^+(x) + \hat{\phi}^-(x) \tag{117}$$

where[8]

$$\hat{\phi}^+(x) \equiv \int \frac{d^3k}{(2\pi)^3}\, \hat{a}(k)\, f_k(x)$$

$$\hat{\phi}^-(x) \equiv \int \frac{d^3k}{(2\pi)^3}\, \hat{a}^\dagger(k)\, f_k^*(x)$$

Then

$$\begin{aligned}
\hat{\phi}(x)\,\hat{\phi}(y) &= \big(\hat{\phi}^+(x) + \hat{\phi}^-(x)\big)\big(\hat{\phi}^+(y) + \hat{\phi}^-(y)\big) \\
&= \hat{\phi}^+(x)\,\hat{\phi}^+(y) + \hat{\phi}^+(x)\,\hat{\phi}^-(y) + \hat{\phi}^-(x)\,\hat{\phi}^+(y) + \hat{\phi}^-(x)\,\hat{\phi}^-(y) \\
&= \hat{\phi}^+(x)\,\hat{\phi}^+(y) + \hat{\phi}^-(y)\,\hat{\phi}^+(x) + \hat{\phi}^-(x)\,\hat{\phi}^+(y) + \hat{\phi}^-(x)\,\hat{\phi}^-(y) \\
&\quad + \big[\hat{\phi}^+(x), \hat{\phi}^-(y)\big]
\end{aligned}$$

In the last step we have commuted one term of operators so that all the $\hat{\phi}^+$ operators are on the right and all the $\hat{\phi}^-$ operators are on the left. We define this ordering of operators

[8] Note that creation operators, $\hat{a}^\dagger(k)$, are in $\hat{\phi}^-$, while the annihilation operators, $\hat{a}(k)$, are in $\hat{\phi}^+$. This notation is confusing, but it is a convention used by many, so we will conform to it.

as the **normal ordered product**:

$$\mathrm{N}\big[\hat{\phi}(x)\,\hat{\phi}(y)\big] \equiv \hat{\phi}^{+}(x)\,\hat{\phi}^{+}(y) + \hat{\phi}^{-}(y)\,\hat{\phi}^{+}(x) + \hat{\phi}^{-}(x)\,\hat{\phi}^{+}(y) + \hat{\phi}^{-}(x)\,\hat{\phi}^{-}(y) \tag{118}$$

The normal ordering of operators has the convenient property that it has zero vacuum expectation value because of the ordering of creation and annihilation operators acting on the vacuum state.

$$\langle 0|\mathrm{N}\big[\hat{\phi}(x)\,\hat{\phi}(y)\big]|0\rangle = 0 \tag{119}$$

Therefore

$$\hat{\phi}(x)\,\hat{\phi}(y) = \mathrm{N}\big[\hat{\phi}(x)\,\hat{\phi}(y)\big] + \big[\hat{\phi}^{+}(x), \hat{\phi}^{-}(y)\big] \tag{120}$$

$$\langle 0|\hat{\phi}(x)\,\hat{\phi}(y)|0\rangle = \cancelto{0}{\langle 0|\mathrm{N}\big[\hat{\phi}(x)\,\hat{\phi}(y)\big]|0\rangle} + \langle 0|\big[\hat{\phi}^{+}(x), \hat{\phi}^{-}(y)\big]|0\rangle$$

Note that because of commutation relation 107,

$$\big[\hat{\phi}^{+}(x), \hat{\phi}^{-}(y)\big] \propto \big[\hat{a}(k), \hat{a}^{\dagger}(k')\big] \propto \hat{\mathbb{1}}$$

$$\Rightarrow \qquad \langle 0|\big[\hat{\phi}^{+}(x), \hat{\phi}^{-}(y)\big]|0\rangle = \big[\hat{\phi}^{+}(x), \hat{\phi}^{-}(y)\big]$$

$$\Rightarrow \qquad \langle 0|\hat{\phi}(x)\,\hat{\phi}(y)|0\rangle = \big[\hat{\phi}^{+}(x), \hat{\phi}^{-}(y)\big]$$

Combining this with 120 gives

$$\mathrm{N}\big[\hat{\phi}(x)\,\hat{\phi}(y)\big] = \hat{\phi}(x)\,\hat{\phi}(y) - \langle 0|\hat{\phi}(x)\,\hat{\phi}(y)|0\rangle \tag{121}$$

Therefore, the normal ordering of operators subtracts the vacuum expectation value of those operators. Therefore, $\hat{P'}_{\mu}$ is the normal ordering of $\hat{P}_{\mu}$.

$$\hat{P'}_{\mu} = \mathrm{N}\big[\hat{P}_{\mu}\big]$$

We will use the operators representing the classical Noether's Charges *normal ordered* as the generators of the Poincaré transformations. Requiring that these operators satisfy the Poincaré algebra will imply the same equal time commutation relations in equations 99 and 100 and these operators will have zero vacuum expectation values.

From now on we will drop the primes from $\hat{P'}_{\mu}$ and $\hat{M'}_{\mu\nu}$ and let $\hat{P}_{\mu}$ and $\hat{M}_{\mu\nu}$ denote normal ordered operators. All this really amounts to is equating $\hat{P}_{\mu}$ to

$$\hat{P}_{\mu} = \int \frac{d^3k}{(2\pi)^3}\, k_{\mu}\, \hat{N}(k) \tag{122}$$

One could argue that we could have just ignored the term creating the zero point energy from the start, but it was worth it to introduce the concept of normal ordering because we will use it in our discussion of Wick's theorem in section 4.2.

3.3.5 Interpretation

We are now in a position to give an interpretation to the eigenstates of the number operators given in equation 113. Consider an example for $n(k)$ given by

$$|n(k)\rangle = \frac{1}{\sqrt{3!}}\, \hat{a}^\dagger(k_1)\, \hat{a}^\dagger(k_2)^3\, |0\rangle \tag{123}$$

Recall that $n(k)$ is an integer valued function of the continuous four-vector k. It denotes the number of times the creation operators $\hat{a}^\dagger(k)$ need to act on the vacuum to give the state $|n(k)\rangle$. Figure 2 illustrates how you might visualize the above example $n(k)$. Equation 122

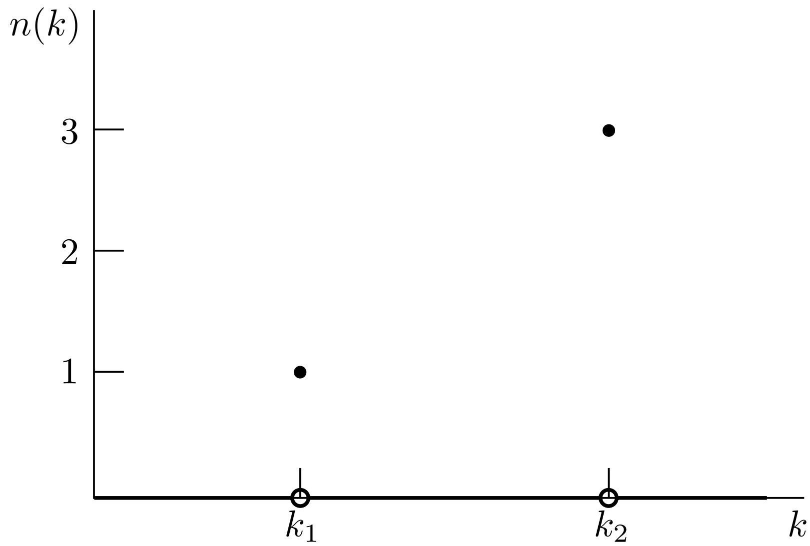


Figure 2: An example $n(k)$ plotted with three of the four-momentum components suppressed.

implies that the four-momentum of this example is

$$P^\mu = k_1^\mu + 3\, k_2^\mu$$

We interpret the state $|n(k)\rangle$ as a state with one particle with four-momentum k_1^μ and three particles with four-momentum k_2^μ. The eigenvalues $n(k)$ of the $\hat{N}(k)$ operators are the number of particles with four-momentum k^μ. As one might have guessed, the creation operator $\hat{a}^\dagger(k)$ *creates* a particle with four-momentum k^μ, and the annihilation operator $\hat{a}(k)$ *destroys* a particle with four-momentum k^μ. If no such particle exists in the state, then

$$\hat{a}(k')\, |n(k)\rangle = 0$$

in accordance with equation 109.

3.3.6 Internal Charge

3.3.7 Spin-statistics

Now we will reface the issue of ambiguity between commutators and anticommutators we encountered in section 3.3.1. One can show that if we had chosen anticommutation relations for the field back in equations 99 and 100, then we would have also derived similar anticommutation relations for the creation and annihilation operators in equations 106 and 107. In particular, we would have shown that

$$\left\{\hat{a}^\dagger(k), \hat{a}^\dagger(k')\right\} = 0 \tag{124}$$

A field that satisfies anticommutation relations, like above, is a **fermion** field. The particles of such a field are fermions. On the other hand, a field that satisfies relations with commutators, like those we derived for the real scalar field, is a **boson** field. The particles this field are bosons.

Note that the state in equation 123 is one example of any integer combinations of creation operators. That is, a state can have any integer number of bosons with same four-momentum. For fermions on the other hand, the anticommutation relation given by equation 124 has an interesting consequence. Note that if the same four-momentum is used as the argument each creation operator, then

$$\left\{\hat{a}^\dagger(k), \hat{a}^\dagger(k)\right\} = \hat{a}^\dagger(k)\,\hat{a}^\dagger(k) + \hat{a}^\dagger(k)\,\hat{a}^\dagger(k) = 0$$

$$\Rightarrow \qquad \hat{a}^\dagger(k)^2 = 0$$

Therefore, one cannot create two of the same fermion with the same four-momentum. This is the explanation of the **Pauli exclusion principle** that says that no two identical fermions may occupy the same quantum state simultaneously. This feature of fermions is the cause of the stability of matter. Without it, electrons and nuclei would not stack into the orbitals and energy levels that make chemistry possible. Instead, they would all fall into the same ground state like bosons in a Bose-Einstein condensate.

Whether or not two identical particles can occupy the same quantum state affects the statistics behind counting the possible states for a system of many particles. This point is of central importance in the study of statistical mechanics.

It turns out that whether a particle is a fermion or boson is directly correlated with its spin. Recall that the real scalar field, a boson field, has spin zero, as we discussed after seeing equation 78. Later on, we will see that the Dirac field has spin 1/2. It will satisfy anticommutation relations, and is therefore a fermion field. This motivates the **spin-statistics theorem**, which states that particles with integer spin are bosons and particles with half integer spin are fermions.

3.3.8 The Feynman Propagator

Recall the plane-wave expansion, equation 103, and let it operate on the vacuum.

$$\begin{aligned} \hat{\phi}(x)\,|0\rangle &= \int \frac{d^3k}{(2\pi)^3} \big(\hat{a}(k)\; f_k(x) + \hat{a}^\dagger(k)\; f_k^*(x)\big)\,|0\rangle \\ &= \int \frac{d^3k}{(2\pi)^3}\, \frac{e^{i\,k\cdot x}}{\sqrt{2\omega_k}}\, \hat{a}^\dagger(k)\,|0\rangle \end{aligned}$$

Note that the state $\hat{\phi}(x)\,|0\rangle$ is the sum of one particle momentum eigenstates. We interpret $\hat{\phi}(x)\,|0\rangle$, not as a state with determined momentum, but as a state with one particle at a *definite point in spacetime*. Similarly,

$$\langle 0|\,\hat{\phi}(y) = \int \frac{d^3k'}{(2\pi)^3}\, \frac{e^{-i\,k'\cdot y}}{\sqrt{2\omega_k'}}\; \langle 0|\,\hat{a}(k')$$

Consider the following inner product.

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

We interpret this as the quantum mechanical amplitude for a particle to be created at spacetime point x and then destroyed at point y, called the "**propagator**". The problem is that this inner product is not Lorentz invariant because it relies on a specific time ordering of events. In order to be physically sensible, the particle must be created at point x *before* it is destroyed at point y. But for spacelike separated points x and y, some observer sees events at x happen *after* y. In order to remedy this, we use the step function defined as

$$\theta(t) \equiv \begin{cases} 0 & \text{if } t < 0 \\ 1 & \text{if } t > 0 \end{cases} \tag{125}$$

Then we define the **time ordered product** as

$$\mathrm{T}\big[\hat{\phi}(y)\;\hat{\phi}(x)\big] \equiv \hat{\phi}(y)\;\hat{\phi}(x)\;\theta(y^0 - x^0) + \hat{\phi}(x)\;\hat{\phi}(y)\;\theta(x^0 - y^0) \tag{126}$$

Basically, the time ordered product orders operators such that those evaluated at later times are to the left. We will see that the correct form of the propagator that is Lorentz invariant is given by

$$\langle 0|\mathrm{T}\big[\hat{\phi}(y)\;\hat{\phi}(x)\big]|0\rangle$$

First consider one term in the time ordering.

$$\begin{aligned}
\langle 0|\hat\phi(y)\,\hat\phi(x)|0\rangle &= \int \frac{d^3k}{(2\pi)^3}\,\frac{d^3k'}{(2\pi)^3}\,\frac{e^{-ik\cdot(y-x)}}{\sqrt{4\,\omega_k\,\omega_{k'}}}\,\langle 0|\hat a(k)\,\hat a^\dagger(k')|0\rangle \\
&= \int \frac{d^3k}{(2\pi)^3}\,\frac{d^3k'}{(2\pi)^3}\,\frac{e^{-ik\cdot(y-x)}}{\sqrt{4\,\omega_k\,\omega_{k'}}}\,\langle 0|\big[\hat a(k),\hat a^\dagger(k')\big]|0\rangle \\
&= \int \frac{d^3k}{(2\pi)^3}\,\frac{d^3k'}{(2\pi)^3}\,\frac{e^{-ik\cdot(y-x)}}{\sqrt{4\,\omega_k\,\omega_{k'}}}\,(2\pi)^3\,\delta^3\left(\vec k-\vec k'\right) \\
&= \int \frac{d^3k}{(2\pi)^3}\,\frac{e^{-ik\cdot(y-x)}}{2\,\omega_k}
\end{aligned}$$

Therefore,

$$\langle 0|\mathrm{T}\big[\hat\phi(y)\,\hat\phi(x)\big]|0\rangle = \int \frac{d^3k}{(2\pi)^3\,2\,\omega_k}\Big(e^{-ik\cdot(y-x)}\,\theta(y^0-x^0)+e^{-ik\cdot(x-y)}\,\theta(x^0-y^0)\Big)$$

The step function has the following integral representation

$$\theta(t) = \frac{-1}{2\pi i}\int_{-\infty}^{\infty} d\omega\,\frac{e^{-i\omega t}}{\omega+i\epsilon} \tag{127}$$

where ϵ is an infinitesimal, positive, real number taken to zero at the end of a calculation. Plugging this in gives,

$$\begin{aligned}
\langle 0|\mathrm{T}\big[\hat\phi(y)\,\hat\phi(x)\big]|0\rangle &= \frac{i}{(2\pi)^4}\int \frac{d^3k\,d\omega}{2\,\omega_k}\left(\frac{e^{-ik\cdot(y-x)-i\omega(y^0-x^0)}}{\omega+i\epsilon}+\frac{e^{+ik\cdot(y-x)+i\omega(y^0-x^0)}}{\omega+i\epsilon}\right) \\
&= \frac{i}{(2\pi)^4}\int \frac{d^3k\,d\omega}{2\,\omega_k}\left(\frac{e^{+i\vec k\cdot(\vec y-\vec x)-i(\omega_k+\omega)(y^0-x^0)}}{\omega+i\epsilon}+\frac{e^{-i\vec k\cdot(\vec y-\vec x)+i(\omega_k+\omega)(y^0-x^0)}}{\omega+i\epsilon}\right)
\end{aligned}$$

In the second term, take $\vec k\to-\vec k$.

$$= \frac{i}{(2\pi)^4}\int \frac{d^3k\,d\omega}{2\,\omega_k}\,e^{+i\vec k\cdot(\vec y-\vec x)}\left(\frac{e^{-i(\omega_k+\omega)(y^0-x^0)}}{\omega+i\epsilon}+\frac{e^{+i(\omega_k+\omega)(y^0-x^0)}}{\omega+i\epsilon}\right)$$

Now in the first term, take $\omega\to\omega-\omega_k$. In the second term, take $\omega\to-\omega-\omega_k$.

$$\begin{aligned}
&= \frac{i}{(2\pi)^4}\int \frac{d^3k\,d\omega}{2\,\omega_k}\,e^{+i\vec k\cdot(\vec y-\vec x)}\,e^{-i\omega(y^0-x^0)}\left(\frac{1}{\omega-\omega_k+i\epsilon}+\frac{1}{-\omega-\omega_k+i\epsilon}\right) \\
&= \frac{i}{(2\pi)^4}\int \frac{d^4k}{2\,\omega_k}\,e^{-ik\cdot(y-x)}\left(\frac{1}{\omega-\omega_k+i\epsilon}-\frac{1}{+\omega+\omega_k-i\epsilon}\right)
\end{aligned}$$

In the last step, we have combined the variables ω and $\vec k$ into a single four-vector, k. Note that their components are unrelated. ω_k is the energy corresponding to momentum $\vec k$ given by $\omega_k=\sqrt{\vec k^2+m^2}$, but omega is free to range over all real values. Cross multiplying the terms gives

$$= \frac{i}{(2\pi)^4} \int \frac{d^4k}{2\,\omega_k}\, e^{-ik\cdot(y-x)} \left(\frac{\cancel{\omega} + \omega_k - i\epsilon - \cancel{\omega} + \omega_k - i\epsilon}{\big(\omega - (\omega_k - i\epsilon)\big)\big(\omega - (-\omega_k + i\epsilon)\big)} \right)$$

We ignore the ϵ terms subtracted from one in the numerator because they are infinitesimal and the only purpose of the ϵ in this expression is to indicate how we should push the poles in the denominator. See Figure ??.

$$= i \int \frac{d^4k}{(2\pi)^4} \frac{e^{-ik\cdot(y-x)}}{\big(\omega - (\omega_k - i\epsilon)\big)\big(\omega - (-\omega_k + i\epsilon)\big)} \tag{128}$$

How we should integrate around these poles was completely dictated by the step function

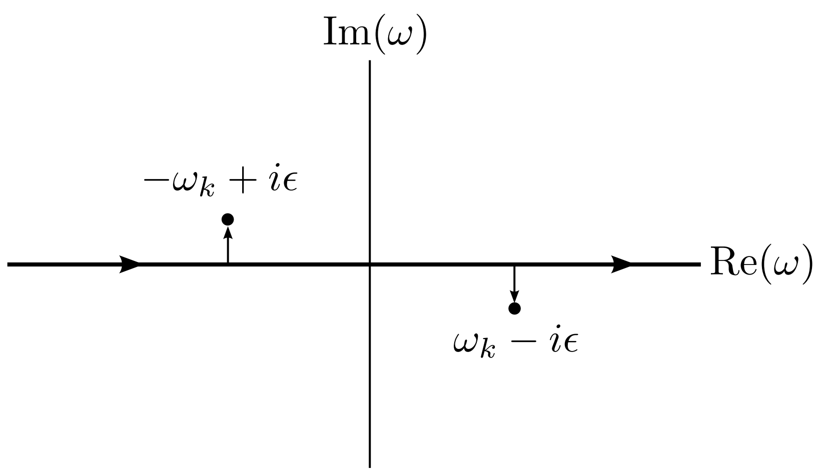


Figure 3: The integration of ω in equation 128 shown in the complex ω plane.

in the time ordering. We arrived at equation 128 in its form so that the orientation of the poles would become evident. In order to simply, we multiple terms back out.

$$\begin{aligned} &= i \int \frac{d^4k}{(2\pi)^4} \frac{e^{-ik\cdot(y-x)}}{\omega^2 - \omega_k^2 + i(-\omega + \omega_k + \omega + \omega_k)\epsilon + \mathcal{O}[\epsilon^2]} \\ &= i \int \frac{d^4k}{(2\pi)^4} \frac{e^{-ik\cdot(y-x)}}{\omega^2 - \vec{k}^2 - m^2 + i\,2\,\omega_k\,\epsilon} \end{aligned}$$

In the last step we have used that $\omega_k^2 = \vec{k}^2 + m^2$. Now take $2\,\omega_k\,\epsilon \to \epsilon$. Therefore,

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

We divide out the i factor and define the **Feynman propagator** as

$$\Delta_F(x-y) \equiv -i\,\langle 0|\mathrm{T}\big[\hat{\phi}(y)\,\hat{\phi}(x)\big]|0\rangle \tag{129}$$

$$\boxed{\Delta_F(x-y) = \int \frac{d^4k}{(2\pi)^4}\, \frac{e^{ik\cdot(x-y)}}{k^2 - m^2 + i\,\epsilon}} \tag{130}$$

While it may not seem obvious at the moment, the propagator is an essential concept in field theory. We will refer back to it numerous times.

We will see that it is often preferable to Fourier transform to get the propagator in momentum space. Evidently, it is given by

$$\begin{aligned}
\Delta_F(k) &= \int d^4x\, e^{-ik\cdot x}\, \Delta_F(x) \\
&= \int d^4x \int \frac{d^4k'}{(2\pi)^4}\, e^{ix\cdot(k'-k)}\, \frac{1}{k'^2 - m^2 + i\,\epsilon} \\
&= \int d^4k'\, \delta(k'-k)\, \frac{1}{k'^2 - m^2 + i\,\epsilon}
\end{aligned}$$

$$\boxed{\Delta_F(k) = \frac{1}{k^2 - m^2 + i\,\epsilon}} \tag{131}$$

4 The Interacting Real Scalar Field

In the free field case, the Lagrangian was simple enough that the equations of motion could be solved and a general solution could be expressed as an expansion in plane-waves. Quantum mechanically, each plane-wave represents a single particle with a certain momentum. In a field configuration that is the sum of several plane-waves, each plane-wave satisfied the equations of motion on its own and therefore evolved in time independent of the other plane-waves. This means that the particles described by the free field do not interact with each other.

We will now add terms to the Lagrangian that will represent interactions. This complication will have the result that the equations of motion can *not* be solved exactly. We will consider the special case of scattering, where particles travel virtually free and only interact briefly in a collision. We will see that the amplitude for a particle to go from x to y can be written as a power series that will converge if the interaction terms are small enough to be thought of as perturbation on the free field Lagrangian.

We will introduce the notation of Feynman diagrams which can be used to pictorally represent the terms of the perturbation series. Finally, we will show that the terms in this perturbation series can be grouped and factored in a way that simplifies calculations immensely.

4.1 Correlation Functions

A Lagrangian for an interacting real scalar field is of the form

$$\mathcal{L} = \mathcal{L}_0 - V$$

where $\mathcal{L}_0$ is the free field Lagrangian given in equation 71, and V, known as the "interaction potential", is a function of the field $\hat{\phi}(x)$. For example, the first interaction potential we will consider is $V(x) = \frac{\lambda}{4!}\,\hat{\phi}^4(x)$, where λ is a real constant.

The presence of V in the Lagrangian density results in the same additional terms in the Hamiltonian density (with the signs flipped[9]).

$$\mathcal{H} = \mathcal{H}_0 + V$$

So the Hamiltonian has the additional terms H_{int}, defined as follows.

$$H = H_0 + H_{\text{int}}$$

$$H_{\text{int}} \equiv \int d^3x\, V$$

[9] recall equation 52

We will come to see that a primary interest in quantum field theory is the calculation of **correlation functions**, also called **Green's functions**. The n-point correlation function, or just the n-point function for short, is defined by

$$G^{(n)}(x_1, x_2, \ldots, x_n) \equiv \langle\Omega|\mathrm{T}\big[\hat{\phi}(x_1)\,\hat{\phi}(x_2)\cdots\hat{\phi}(x_n)\big]|\Omega\rangle \tag{132}$$

where $|\Omega\rangle$ is the vacuum state of the interacting field. It will be our goal to calculate the two point Green's function, $\langle\Omega|\mathrm{T}\big[\hat{\phi}(x_1)\,\hat{\phi}(x_2)\big]|\Omega\rangle$, which analogous to the Feynman propagator of the free field theory, represents the amplitude for a particle to propagate from x_1 to x_2. The interaction will affect this calculation in two ways: the interaction terms in the Hamiltonian, and the distinction of $|\Omega\rangle$ from the vacuum of free field, $|0\rangle$.

Like any Heisenberg operator, the field evolves as

$$\hat{\phi}(t, \vec{x}) = e^{i\hat{H}\,(t-t_0)}\,\hat{\phi}(t_0, \vec{x})\,e^{-i\hat{H}\,(t-t_0)}$$

As we turn off the interaction, $V \to 0$ and $H \to H_0$, then

$$\hat{\phi}(t, \vec{x}) \to \hat{\phi}_I(t, \vec{x}) \equiv e^{i\hat{H}_0\,(t-t_0)}\,\hat{\phi}(t_0, \vec{x})\,e^{-i\hat{H}_0\,(t-t_0)} \tag{133}$$

$\hat{\phi}_I$ is called the "interaction picture field," or just the "interaction field." The interaction picture is an intermediate picture between the Schrödinger picture and the Heisenberg picture, where the time dependence of the free Hamiltonian is carried by the operators, but the time dependence due to the interaction part of the Hamiltonian is carried by the state vectors.

Since $\hat{\phi}_I(t, \vec{x}) = \hat{\phi}(t, \vec{x})\Big|_{V=0}$, then $\hat{\phi}_I$ satisfies the same equation of motion as the free real scalar field, the Klein-Gordon equation 102. Inverting the definition of the interaction field, equation 133, implies that

$$\hat{\phi}(t_0, \vec{x}) = e^{-i\hat{H}_0\,(t-t_0)}\,\hat{\phi}_I(t, \vec{x})\,e^{i\hat{H}_0\,(t-t_0)} \tag{134}$$

So for general times

$$\begin{aligned}\hat{\phi}(t, \vec{x}) &= e^{+i\hat{H}\,(t-t_0)}\,e^{-i\hat{H}_0\,(t-t_0)}\,\hat{\phi}_I(t, \vec{x})\,e^{+i\hat{H}_0\,(t-t_0)}\,e^{-i\hat{H}\,(t-t_0)} \\ &= \hat{U}^\dagger(t, t_0)\,\hat{\phi}_I(t, \vec{x})\,\hat{U}(t, t_0)\end{aligned} \tag{135}$$

where we have defined the "interaction picture time evolution operator," $\hat{U}(t, t_0)$, as

$$\hat{U}(t_1, t_2) \equiv e^{+i\hat{H}_0\,(t_1-t_2)}\,e^{-i\hat{H}\,(t_1-t_2)}$$

Explicitly taking the time derivative of $\hat{U}(t,t_0)$ gives

$$\begin{aligned}\frac{\partial}{\partial t}\,\hat{U}(t,t_0) &= e^{+i\hat{H}_0\,(t-t_0)}(i\hat{H}_0 - i\hat{H})e^{-i\hat{H}\,(t-t_0)} \\ &= -i\,e^{+i\hat{H}_0\,(t-t_0)}\,\hat{H}_{\text{int}}\,e^{-i\hat{H}\,(t-t_0)} \\ &= -i\,e^{+i\hat{H}_0\,(t-t_0)}\,\hat{H}_{\text{int}}\,e^{-i\hat{H}_0\,(t-t_0)}\,e^{+i\hat{H}_0\,(t-t_0)}\,e^{-i\hat{H}\,(t-t_0)}\end{aligned}$$

In this last step we have inserted a factor of $1 = e^{-i\hat{H}_0\,(t-t_0)}\,e^{+i\hat{H}_0\,(t-t_0)}$. Now we define the interaction picture interaction Hamiltonian:

$$\hat{H}_I(t) \equiv e^{+i\hat{H}_0\,(t-t_0)}\,\hat{H}_{\text{int}}(t)\,e^{-i\hat{H}_0\,(t-t_0)}$$

Then we see that $\hat{U}(t,t_0)$ satisfies a Schrödinger-like equation:

$$i\,\frac{\partial}{\partial t}\,\hat{U}(t,t_0) = \hat{H}_I(t)\,\hat{U}(t,t_0) \tag{136}$$

If $\hat{U}(t,t_0)$ and $\hat{H}_I(t)$ were just functions and not operators, then we could divide both sides by U and integrate getting the simple solution of $U(t,t_0) = \exp\left(-i\int_{t_0}^{t} dt'\,H(t')\right)$. But because $\hat{U}(t,t_0)$ is an operator, we do this more carefully, treating $\hat{U}(t,t_0)$ as an operator and noting its ordering. Integrating the differential equation 136, from t_0 to t gives.

$$\hat{U}(t_1,t_0)\Big|_{t_1=t_0}^{t} = -i\int_{t_0}^{t} dt_1\,\hat{H}_I(t_1)\,\hat{U}(t_1,t_0)$$

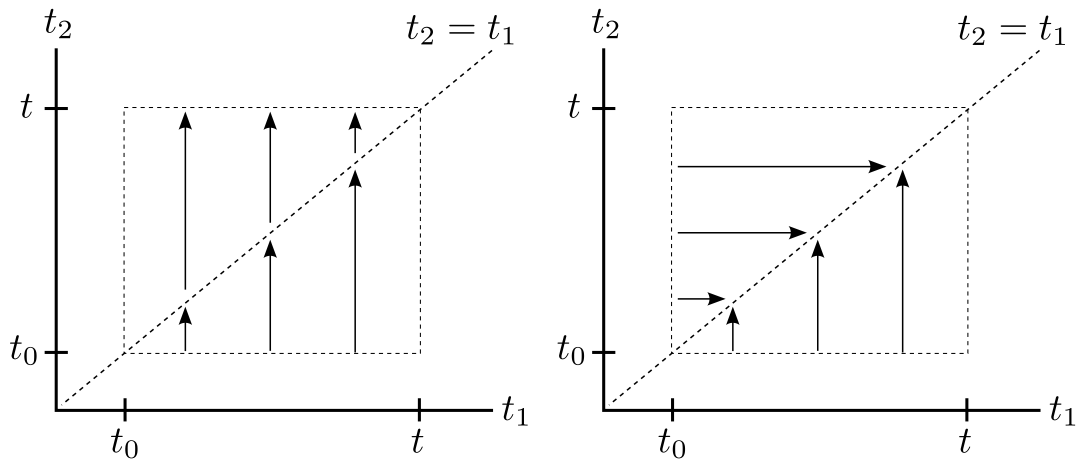
We now use the initial condition that $\hat{U}(t_0,t_0) = \hat{\mathbb{1}}$ because we know from equations 135 that when $t = t_0$, it should be that $\hat{\phi}(t,\vec{x}) = \hat{\phi}_I(t,\vec{x})$.

$$\hat{U}(t,t_0) = \hat{\mathbb{1}} - i\int_{t_0}^{t} dt_1\,\hat{H}_I(t_1)\,\hat{U}(t_1,t_0)$$

This "solution" only gives $\hat{U}(t,t_0)$ in terms of an integral of itself. Iteratively plugging this expression into itself gives the solution known as the "Dyson series."

$$\begin{aligned}\hat{U}(t,t_0) &= \hat{\mathbb{1}} + (-i)\int_{t_0}^{t} dt_1\,\hat{H}_I(t_1) + (-i)^2\int_{t_0}^{t} dt_1\int_{t_0}^{t_1} dt_2\,\hat{H}_I(t_1)\,\hat{H}_I(t_2) \\ &\quad + (-i)^3\int_{t_0}^{t} dt_1\int_{t_0}^{t_1} dt_2\int_{t_0}^{t_2} dt_3\,\hat{H}_I(t_1)\,\hat{H}_I(t_2)\,\hat{H}_I(t_3) + \cdots\end{aligned} \tag{137}$$

Note that the integrals are nested in the sense that the limits of the inner integrals depend on the integration variables of the outer integrals. Also note that because the limits of integration enforce that $t_1 \geq t_2 \geq \cdots \geq t_{n-1} \geq t_n \geq t_0$, the $\hat{H}_I$ operators are all time ordered. We can un-nest these integrals with the following trick. Consider the second order



(a) equation 138 (b) equation 139

Figure 4: Un-nesting integrals

term for example.

$$I \equiv \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2\, \hat{H}_I(t_1)\, \hat{H}_I(t_2)$$

Now consider the following expression without nested integrals,

$$I' \equiv \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2\, \mathrm{T}\big[\hat{H}_I(t_1)\, \hat{H}_I(t_2)\big]$$

which we will relate to the previous integrals by the following

$$I' = \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2\, \hat{H}_I(t_1)\, \hat{H}_I(t_2) + \int_{t_0}^{t} dt_1 \int_{t_1}^{t} dt_2\, \hat{H}_I(t_2)\, \hat{H}_I(t_1) \tag{138}$$

$$= I + \int_{t_0}^{t} dt_2 \int_{t_0}^{t_2} dt_1\, \hat{H}_I(t_2)\, \hat{H}_I(t_1) \tag{139}$$

$$= I + \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2\, \hat{H}_I(t_1)\, \hat{H}_I(t_2) \tag{140}$$

$$= 2\, I$$

In equation 139, we have changed the limits of integration from those depicted in figure 4(a) to those depicted in 4(b), while still integrating over the same region. In equation 140, we have simply switched the names of the dummy variables $t_1 \leftrightarrow t_2$. Therefore,

$$\int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2\, \hat{H}_I(t_1)\, \hat{H}_I(t_2) = \frac{1}{2} \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2\, \mathrm{T}\big[\hat{H}_I(t_1)\, \hat{H}_I(t_2)\big]$$

and in general,

$$\int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{n-1}} dt_n\, \hat{H}_I(t_1) \cdots \hat{H}_I(t_n) = \frac{1}{n!} \int_{t_0}^{t} dt_1 \cdots dt_n\, \mathrm{T}\big[\hat{H}_I(t_1) \cdots \hat{H}_I(t_n)\big]$$

Plugging this result into the Dyson series, equation 137, gives

$$\begin{aligned} \hat{U}(t, t_0) &= \hat{\mathbb{1}} + \frac{(-i)}{1!} \int_{t_0}^{t} dt_1\, \hat{H}_I(t_1) + \frac{(-i)^2}{2!} \int_{t_0}^{t} dt_1\, dt_2\, \mathrm{T}\big[\hat{H}_I(t_1)\, \hat{H}_I(t_2)\big] \\ &\quad + \frac{(-i)^3}{3!} \int_{t_0}^{t} dt_1\, dt_2\, dt_3\, \mathrm{T}\big[\hat{H}_I(t_1)\, \hat{H}_I(t_2)\, \hat{H}_I(t_3)\big] + \cdots \end{aligned}$$

which we will write compactly as

$$\hat{U}(t, t_0) = \mathrm{T}\left[\exp\left(-i \int_{t_0}^{t} dt'\, \hat{H}_I(t')\right)\right] \tag{141}$$

From the properties of limits of integrals it is now evident that adjacent $\hat{U}$ operators, with the appropriate arguments, can be combined into one[10].

$$\begin{aligned} \hat{U}(t_1, t_2)\, \hat{U}(t_2, t_3) &= \mathrm{T}\left[\exp\left(-i \int_{t_2}^{t_1} dt'\, \hat{H}_I(t') - i \int_{t_3}^{t_2} dt''\, \hat{H}_I(t'')\right)\right] \\ &= \mathrm{T}\left[\exp\left(-i \int_{t_3}^{t_1} dt'\, \hat{H}_I(t')\right)\right] \end{aligned}$$

$$\hat{U}(t_1, t_2)\, \hat{U}(t_2, t_3) = \hat{U}(t_1, t_3) \tag{142}$$

$$\begin{aligned} \hat{U}^\dagger(t_2, t_1) &= \mathrm{T}\left[\exp\left(-i \int_{t_1}^{t_2} dt'\, \hat{H}_I(t')\right)^\dagger\right] \\ &= \mathrm{T}\left[\exp\left(+i \int_{t_1}^{t_2} dt'\, \hat{H}_I(t')\right)\right] \\ &= \mathrm{T}\left[\exp\left(-i \int_{t_2}^{t_1} dt'\, \hat{H}_I(t')\right)\right] \end{aligned}$$

$$\hat{U}^\dagger(t_2, t_1) = \hat{U}(t_1, t_2) \tag{143}$$

Therefore,

$$\hat{U}(t_1, t_2)\, \hat{U}^\dagger(t_3, t_2) = \hat{U}(t_1, t_3) \tag{144}$$

We are almost ready to derive what will be a very useful relation for correlation functions,

[10] Note that in combining the exponentials, we are not complicated by the Baker-Campbell-Hausdorff formula (shown below) because thankfully, $\hat{H}_I(t')$ commutes with $\hat{H}_I(t'')$ for all times, a fact most readily seen because we have required that the Poincaré algebra be satisfied (see equation 92). In general, one has to be careful when combining exponentials of operators. $e^{(\hat{A}+\hat{B})} = e^{\hat{A}}\, e^{\hat{B}}\, e^{-\frac{1}{2}[\hat{A},\hat{B}]}\, e^{\frac{1}{6}\left(2[\hat{B},[\hat{A},\hat{B}]] + [\hat{A},[\hat{A},\hat{B}]]\right)} \ldots$

but first we turn our attention to the vacuum, $|\Omega\rangle$. We will relate the vacuum state of the interacting field, $|\Omega\rangle$, to the vacuum of a free field, $|0\rangle$, by expanding $|0\rangle$ in terms of the eigenstates of the complete Hamiltonian with interaction terms, and then evolving it in time with the complete Hamiltonian.

$$|0\rangle = \sum_n |n\rangle \langle n|0\rangle$$

where $|n\rangle$ are the eigenstates of $\hat{H}$.

$$\hat{H}\,|n\rangle = E_n\,|n\rangle$$

Now evolve it in time, but with the complete interacting Hamiltonian.

$$e^{-i\,\hat{H}\,t}\,|0\rangle = \sum_n e^{-i\,E_n\,t}\,|n\rangle \langle n|0\rangle$$

Since the ground state (vacuum) is a term in this sum, we can pull it out explicitly,

$$e^{-i\,\hat{H}\,t}\,|0\rangle = e^{-i\,E_0\,t}\,|\Omega\rangle \langle \Omega|0\rangle + \sum_{n\neq 0} e^{-i\,E_n\,t}\,|n\rangle \langle n|0\rangle \tag{145}$$

where $E_0 = \langle \Omega|\hat{H}|\Omega\rangle$ is the ground state energy of the interacting theory. We define zero energy to be the energy of the ground state of the noninteracting theory, $\hat{H}_0\,|0\rangle = 0$. Now, in order to kill off all the other terms in the sum, we take time to infinity in a direction slightly rotated towards the negative imaginary.

$$\lim_{t\to\infty(1-i\,\epsilon)} e^{-i\,\hat{H}\,t}\,|0\rangle = e^{-i\,E_0\,t}\,|\Omega\rangle \langle \Omega|0\rangle + \sum_{n\neq 0} \cancelto{0}{e^{-i\,E_n\,t}}\,|n\rangle \langle n|0\rangle$$

These terms are subdominant to the term with E_0 because E_0 is the smallest of all E_n. Inverting this expression for $|\Omega\rangle$ gives

$$|\Omega\rangle = \lim_{t\to\infty(1-i\,\epsilon)} \left(e^{-i\,E_0\,t}\,\langle \Omega|0\rangle\right)^{-1} e^{-i\,\hat{H}\,t}\,|0\rangle$$

Since t is now very large, we can shift it by a finite constant, t_0.

$$|\Omega\rangle = \lim_{t\to\infty(1-i\,\epsilon)} \left(e^{-i\,E_0\,(t+t_0)}\,\langle \Omega|0\rangle\right)^{-1} e^{-i\,\hat{H}\,(t+t_0)}\,|0\rangle$$

And we can insert a factor of $e^{-i\,\hat{H}_0(-t-t_0)}$ because recall that $\hat{H}_0\,|0\rangle = 0$.

$$
\begin{aligned}
|\Omega\rangle &= \lim_{t\to\infty(1-i\,\epsilon)} \left(e^{-i\,E_0\,(t+t_0)}\,\langle\Omega|0\rangle\right)^{-1} \underbrace{e^{+i\,\hat{H}\,((-t)-t_0)}\,e^{-i\,\hat{H}_0((-t)-t_0)}}_{\hat{U}^\dagger(-t,t_0)}|0\rangle \\
&= \lim_{t\to\infty(1-i\,\epsilon)} \left(e^{-i\,E_0\,(t+t_0)}\,\langle\Omega|0\rangle\right)^{-1} \hat{U}(t_0,-t)\,|0\rangle \qquad (146)
\end{aligned}
$$

Note that this indicates that one can get from the free vacuum state to the interacting vacuum by evolving it with the interaction picture time evolution operator from the asymtotically infinite past to the reference time t_0. The rest is just a normalization factor.

Similar to equation 145, we can express $\langle\Omega|$ as follows.

$$
\langle\Omega|\,e^{-i\,\hat{H}\,t} = \langle 0|\Omega\rangle\,\langle\Omega|\,e^{-i\,E_0\,t} + \sum_{n\neq 0}\langle 0|n\rangle\,\langle n|\,e^{-i\,E_n\,t}
$$

Following the same reasoning we have

$$
\langle\Omega| = \lim_{t\to\infty(1-i\,\epsilon)} \left(e^{-i\,E_0\,t}\,\langle 0|\Omega\rangle\right)^{-1}\langle 0|\,e^{-i\,\hat{H}\,t}
$$

Shifting $t \to t - t_0$ and inserting a factor of $e^{+i\,\hat{H}_0\,(t-t_0)}$ gives

$$
\begin{aligned}
\langle\Omega| &= \lim_{t\to\infty(1-i\,\epsilon)} \left(e^{-i\,E_0\,(t-t_0)}\,\langle 0|\Omega\rangle\right)^{-1}\langle 0|\,e^{+i\,\hat{H}_0\,(t-t_0)}e^{-i\,\hat{H}\,(t-t_0)} \\
&= \lim_{t\to\infty(1-i\,\epsilon)} \left(e^{-i\,E_0\,(t-t_0)}\,\langle 0|\Omega\rangle\right)^{-1}\langle 0|\,\hat{U}(t,t_0) \qquad (147)
\end{aligned}
$$

Combining our results, we can now write the 2-point function in terms of the free vacuum and interaction fields. For the moment, assume $x_1^0 > x_2^0 > t_0$ so that we can explicitly write out the time ordering. Then we use equations 132, 134, 146, and 147.

$$
\langle\Omega|\mathrm{T}\big[\hat{\phi}(x_1)\,\hat{\phi}(x_2)\big]|\Omega\rangle = \langle\Omega|\hat{\phi}(x_1)\,\hat{\phi}(x_2)|\Omega\rangle
$$

$$
\begin{aligned}
&= \lim_{t\to\infty(1-i\,\epsilon)} \left(e^{-i\,2\,E_0\,t)}\,|\langle 0|\Omega\rangle|^2\right)^{-1} \\
&\quad \langle 0|\hat{U}(t,t_0)\,\hat{U}^\dagger(x_1^0,t_0)\,\hat{\phi}_I(x_1)\,\hat{U}(x_1^0,t_0)\,\hat{U}^\dagger(x_2^0,t_0)\,\hat{\phi}_I(x_2)\,\hat{U}(x_2^0,t_0)\,\hat{U}(t_0,-t)|0\rangle \\
&= \lim_{t\to\infty(1-i\,\epsilon)} \left(e^{-i\,2\,E_0\,t)}\,|\langle 0|\Omega\rangle|^2\right)^{-1}\langle 0|\hat{U}(t,x_1^0)\,\hat{\phi}_I(x_1)\,\hat{U}(x_1^0,x_2^0)\,\hat{\phi}_I(x_2)\,\hat{U}(x_2^0,-t)|0\rangle
\end{aligned}
$$

We have used the property of $\hat{U}$ operators that allows one to combine them, equation 144.

We can get rid of the awkward factor in front by dividing by a convenient form of 1:

$$\begin{aligned} 1 &= \langle\Omega|\Omega\rangle = \lim_{t\to\infty(1-i\,\epsilon)} \left(e^{-i\,2\,E_0\,t)}\,|\langle 0|\Omega\rangle|^2\right)^{-1} \langle 0|\hat{U}(t,t_0)\,\hat{U}(t_0,-t)|0\rangle \\ &= \lim_{t\to\infty(1-i\,\epsilon)} \left(e^{-i\,2\,E_0\,t)}\,|\langle 0|\Omega\rangle|^2\right)^{-1} \langle 0|\hat{U}(t,-t)|0\rangle \end{aligned}$$

Therefore,

$$\begin{aligned} \langle\Omega|\mathrm{T}\big[\hat{\phi}(x_2)\,\hat{\phi}(x_1)\big]|\Omega\rangle &= \lim_{t\to\infty(1-i\,\epsilon)} \frac{\langle 0|\hat{U}(t,x_1^0)\,\hat{\phi}_I(x_1)\,\hat{U}(x_1^0,x_2^0)\,\hat{\phi}_I(x_2)\,\hat{U}(x_2^0,-t)|0\rangle}{\langle 0|\hat{U}(t,-t)|0\rangle} \\ &= \lim_{t\to\infty(1-i\,\epsilon)} \frac{\langle 0|\mathrm{T}\big[\hat{\phi}_I(x_1)\,\hat{\phi}_I(x_2)\,\hat{U}(t,x_1^0)\,\hat{U}(x_1^0,x_2^0)\,\hat{U}(x_2^0,-t)\big]|0\rangle}{\langle 0|\hat{U}(t,-t)|0\rangle} \\ &= \lim_{t\to\infty(1-i\,\epsilon)} \frac{\langle 0|\mathrm{T}\big[\hat{\phi}_I(x_1)\,\hat{\phi}_I(x_2)\,\hat{U}(t,-t)\big]|0\rangle}{\langle 0|\hat{U}(t,-t)|0\rangle} \end{aligned}$$

And finally, plugging in our integral solution for $\hat{U}(t,-t)$, equation 141 gives our formula for the 2-point correlation function.

$$\boxed{\langle\Omega|\mathrm{T}\big[\hat{\phi}(x_1)\,\hat{\phi}(x_2)\big]|\Omega\rangle = \lim_{t\to\infty(1-i\,\epsilon)} \frac{\langle 0|\mathrm{T}\left[\hat{\phi}_I(x_1)\,\hat{\phi}_I(x_2)\;\exp\left(-i\int_{-t}^{t} dt'\;\hat{H}_I(t')\right)\right]|0\rangle}{\langle 0|\mathrm{T}\left[\exp\left(-i\int_{-t}^{t} dt'\;\hat{H}_I(t')\right)\right]|0\rangle}} \tag{148}$$

It may not seem like we have made much progress, but we actually have. We will see that the expression on the right is one that we can actually calculate because it is in terms of the free vacuum and the interaction picture fields, which have the same equation of motion as a free field. Because of this, we will see that we can calculate these expectation values much in the same way we were able to calculate the expectation value called the Feynman propagator for free field theory in Section 3.3.8. Also, we will see that the fact that this expression is the ratio of expectation values will allow many terms to factor and cancel.

For the general n-point function, our derivation of the general formula analogous to equation 148 would be exactly the same except we would have had more $\hat{\phi}_I$ operators to write down and more $\hat{U}$ operators to lump together. The general formula is

$$\begin{aligned} &\langle\Omega|\mathrm{T}\big[\hat{\phi}(x_1)\,\hat{\phi}(x_2)\cdots\hat{\phi}(x_n)\big]|\Omega\rangle = \\ &\lim_{t\to\infty(1-i\,\epsilon)} \frac{\langle 0|\mathrm{T}\left[\hat{\phi}_I(x_1)\,\hat{\phi}_I(x_2)\cdots\hat{\phi}_I(x_n)\;\exp\left(-i\int_{-t}^{t} dt'\;\hat{H}_I(t')\right)\right]|0\rangle}{\langle 0|\mathrm{T}\left[\exp\left(-i\int_{-t}^{t} dt'\;\hat{H}_I(t')\right)\right]|0\rangle} \end{aligned} \tag{149}$$

To make things more concrete let's choose a specfic $\hat{H}_I$ to plug into equation 148. It is called the ϕ^4 theory. Let

$$\hat{V}(x) = \frac{\lambda}{4!}\,\hat{\phi}^4(x)$$

where λ is a small positive real number. Then

$$\hat{H}_I = \int d^3x\,\frac{\lambda}{4!}\,\hat{\phi}_I^4(x)$$

Equation 148 is suited for doing purturbative calculations because since λ is small, the Taylor expansion of the exponential is a purturbation series in λ. The first few terms of the expansion of the numerator in equation 148 are

$$\begin{aligned}
&\langle 0|\mathrm{T}\left[\hat{\phi}_I(x_1)\,\hat{\phi}_I(x_2)\,\exp\left(-i\int_{-t}^{t} dt'\,\hat{H}_I(t')\right)\right]|0\rangle = \\
&\quad \langle 0|\mathrm{T}\big[\hat{\phi}_I(x_1)\,\hat{\phi}_I(x_2)\big]|0\rangle \\
&+\ \frac{1}{1!}\left(\frac{-i\,\lambda}{4!}\right)\int d^4y_1\,\langle 0|\mathrm{T}\left[\hat{\phi}_I(x_1)\,\hat{\phi}_I(x_2)\,\hat{\phi}_I^4(y_1)\right]|0\rangle \\
&+\ \frac{1}{2!}\left(\frac{-i\,\lambda}{4!}\right)^2\int d^4y_1\,d^4y_2\,\langle 0|\mathrm{T}\left[\hat{\phi}_I(x_1)\,\hat{\phi}_I(x_2)\,\hat{\phi}_I^4(y_1)\,\hat{\phi}_I^4(y_2)\right]|0\rangle \\
&+\ \cdots
\end{aligned} \tag{150}$$

We have combined the time integral from equation 148 with the space integral from $\hat{H}_I$ into an integral overall all spacetime with the caviet that the time integral is infinitesmially rotated towards the imaginary by the $t \to \infty(1 - i\,\epsilon)$ limit. Each successive term in this series is supressed by higher powers of λ, therefore we only need to calculate as many terms as our precision requires or until the calculation gets unbearably complicated.

4.2 Wick's Theorem

What we have really accomplished in equation 148 is to relate expressions like $\langle\Omega|\mathrm{T}\big[\hat{\phi}(x_2)\,\hat{\phi}(x_1)\big]|\Omega\rangle$ to a series of expressions like $\langle 0|\mathrm{T}\big[\hat{\phi}_I(x_1)\,\hat{\phi}_I(x_2)\,\hat{\phi}_I(x_3)\cdots\big]|0\rangle$. In this section, we will derive a theorem called Wick's theorem that allows us to write any expression of the form of $\langle 0|\mathrm{T}\big[\hat{\phi}_I(x_1)\,\hat{\phi}_I(x_2)\,\hat{\phi}_I(x_3)\cdots\big]|0\rangle$ as a product of 2-point functions like $\langle 0|\mathrm{T}\big[\hat{\phi}_I(x_1)\,\hat{\phi}_I(x_2)\big]|0\rangle$.

Consider the following time-ordered product, where we know $y^0 > x^0$.

$$\begin{aligned}
\mathrm{T}\big[\hat{\phi}(y)\,\hat{\phi}(x)\big] &= \hat{\phi}(y)\,\hat{\phi}(x) \\
&= \big(\hat{\phi}^+(y) + \hat{\phi}^-(y)\big)\big(\hat{\phi}^+(x) + \hat{\phi}^-(x)\big) \\
&= \hat{\phi}^+(y)\,\hat{\phi}^+(x) + \hat{\phi}^+(y)\,\hat{\phi}^-(x) + \hat{\phi}^-(y)\,\hat{\phi}^+(x) + \hat{\phi}^-(y)\,\hat{\phi}^-(x) \\
&= \hat{\phi}^+(y)\,\hat{\phi}^+(x) + \hat{\phi}^-(x)\,\hat{\phi}^+(y) + \hat{\phi}^-(y)\,\hat{\phi}^+(x) + \hat{\phi}^-(y)\,\hat{\phi}^-(x) \\
&\quad + \big[\hat{\phi}^+(y), \hat{\phi}^-(x)\big] \\
&= \mathrm{N}\big[\hat{\phi}(y)\,\hat{\phi}(x)\big] + \big[\hat{\phi}^+(y), \hat{\phi}^-(x)\big]
\end{aligned}$$

We are using the $\hat{\phi}^{\pm}$ notation we introduced in equation 117. In the last step we have commuted one term of operators so that all the $\hat{\phi}^{+}$ operators are on the right and all the $\hat{\phi}^{-}$ operators are on the left. This is what we defined as the normal-ordered product in equation 118.

For general times we have

$$\mathrm{T}\big[\hat{\phi}(y)\,\hat{\phi}(x)\big] = \mathrm{N}\big[\hat{\phi}(y)\,\hat{\phi}(x)\big] + \begin{cases} \big[\hat{\phi}^{+}(y), \hat{\phi}^{-}(x)\big] & \text{if } y^0 > x^0 \\ \big[\hat{\phi}^{+}(x), \hat{\phi}^{-}(y)\big] & \text{if } x^0 > y^0 \end{cases} \tag{151}$$

We will define the **Wick contraction** of two field operators as follows to denote this difference between time ordering and normal ordering.

$$\wick{\c{\hat{\phi}}(y)\c{\hat{\phi}}(x)} \equiv \mathrm{T}\big[\hat{\phi}(y)\,\hat{\phi}(x)\big] - \mathrm{N}\big[\hat{\phi}(y)\,\hat{\phi}(x)\big] \tag{152}$$

Now we state an identity, known as **Wick's theorem**, for converting time order products of multiple operators to a sum of normal ordered products and contractions.

$$\mathrm{T}\big[\hat{\phi}(x_1)\,\hat{\phi}(x_2)\cdots\hat{\phi}(x_n)\big] = \mathrm{N}\big[\hat{\phi}(x_1)\,\hat{\phi}(x_2)\cdots\hat{\phi}(x_n)\big] \;+\; \text{all possible contractions} \tag{153}$$

To see explicitly what is meant by "all possible contractions," consider the case of $n = 4$. For brevity, let ϕ_i denote $\hat{\phi}(x_i)$.

$$\begin{aligned} \mathrm{T}\big[\phi_1\,\phi_2\,\phi_3\,\phi_4\big] \;=\;& \mathrm{N}\big[\phi_1\,\phi_2\,\phi_3\,\phi_4\big] + \wick{\c\phi_1\,\c\phi_2}\,\phi_3\,\phi_4 + \wick{\c\phi_1\,\phi_2\,\c\phi_3}\,\phi_4 + \wick{\c\phi_1\,\phi_2\,\phi_3\,\c\phi_4} \\ & + \phi_1\,\wick{\c\phi_2\,\c\phi_3}\,\phi_4 + \phi_1\,\wick{\c\phi_2\,\phi_3\,\c\phi_4} + \phi_1\,\phi_2\,\wick{\c\phi_3\,\c\phi_4} \\ & + \wick{\c\phi_1\,\c\phi_2\,\c\phi_3\,\c\phi_4} + \wick{\c1\phi_1\,\c2\phi_2\,\c1\phi_3\,\c2\phi_4} + \wick{\c1\phi_1\,\c2\phi_2\,\c2\phi_3\,\c1\phi_4} \end{aligned}$$

The right-hand side of this identity is known as the "Wick expansion." When a term has uncontracted fields, our notation means that those uncontracted fields are normal ordered.

$$\wick{\c\phi_1\,\phi_2\,\c\phi_3}\,\phi_4 \equiv \wick{\c\phi_1\,\c\phi_3}\,\mathrm{N}\big[\phi_3\,\phi_4\big]$$

To prove Wick's theorem, one uses induction. Wick's theorem is true for $n = 2$ by definition. We assume it is true for $n-1$ and prove that it is true for n. Let us label x_i

such that $x_1^0 \geq x_2^0 \geq \cdots \geq x_n^0$.

$$\begin{aligned}
\mathrm{T}\big[\phi_1\,\phi_2\cdots\phi_n\big] &= \phi_1\,\phi_2\cdots\phi_n \\
&= \phi_1 \left(\mathrm{N}\big[\phi_2\cdots\phi_n\big] + \begin{pmatrix}\text{all contractions}\\ \text{not involving } \phi_1\end{pmatrix}\right) \\
&= (\phi_1^+ + \phi_1^-) \left(\mathrm{N}\big[\phi_2\cdots\phi_n\big] + \begin{pmatrix}\text{all contractions}\\ \text{not involving } \phi_1\end{pmatrix}\right)
\end{aligned}$$

In the step before last, we used Wick's theorem for $\phi_2 \cdots \phi_n$. To make this last step look like Wick's theorem, we want to pull the ϕ_1 into the normal-ordered product. The ϕ_1^- is already normal-ordered so we can just bring it in, but the ϕ_1^+ must be commuted passed all the other operators.

$$\begin{aligned}
\phi_1^+\,\mathrm{N}\big[\phi_2\cdots\phi_n\big] &= \mathrm{N}\big[\phi_2\cdots\phi_n\big]\phi_1^+ + \big[\phi_1^+, \mathrm{N}\big[\phi_2\cdots\phi_n\big]\big] \\
&= \mathrm{N}\big[\phi_1^+\,\phi_2\cdots\phi_n\big] + \big[\phi_1^+,\phi_2^-\big]\,\mathrm{N}\big[\phi_3\cdots\phi_n\big] + \big[\phi_1^+,\phi_3^-\big]\mathrm{N}\big[\phi_2\,\phi_3\cdots\phi_n\big] + \cdots \\
&= \mathrm{N}\big[\phi_1^+\,\phi_2\cdots\phi_n\big] + \big[\phi_1^+,\phi_2^-\big]\,\mathrm{N}\big[\phi_3\cdots\phi_n\big] + \big[\phi_1^+,\phi_3^-\big]\mathrm{N}\big[\phi_2\,\phi_3\cdots\phi_n\big] + \cdots \\
&= \mathrm{N}\big[\phi_1^+\,\phi_2\,\phi_3\,\phi_4\big] + \contraction{}{\phi_1}{}{\phi_2}\phi_1\,\phi_2\,\phi_3\,\phi_4 + \contraction{}{\phi_1}{\,\phi_2\,}{\phi_3}\phi_1\,\phi_2\,\phi_3\,\phi_4 + \cdots \\
&= \mathrm{N}\big[\phi_1^+\,\phi_2\,\phi_3\,\phi_4\big] + \begin{pmatrix}\text{all contractions}\\ \text{involving } \phi_1\end{pmatrix}
\end{aligned}$$

Combining this with the term multiplied by ϕ_1^- gives back the form of Wick's theorem.

$$\begin{aligned}
\mathrm{T}\big[\phi_1\,\phi_2\cdots\phi_n\big] &= \mathrm{N}\big[\phi_1^+\,\phi_2\cdots\phi_4\big] + \begin{pmatrix}\text{all contractions}\\ \text{involving } \phi_1\end{pmatrix} \\
&\quad + \phi_1^- \left(\mathrm{N}\big[\phi_2\cdots\phi_n\big] + \begin{pmatrix}\text{all contractions}\\ \text{not involving } \phi_1\end{pmatrix}\right) \\
&= \mathrm{N}\big[\phi_1\,\phi_2\cdots\phi_n\big] + \text{all possible contractions}
\end{aligned}$$

This completes our proof of Wick's theorem.

Now recall that what we are really interested in calculating is the free vacuum expectation value of the time ordered product of interaction picture fields, $\langle 0|\mathrm{T}\big[\hat{\phi}_I(x_1)\,\hat{\phi}_I(x_2)\,\hat{\phi}_I(x_3)\cdots\big]|0\rangle$. Using Wick's theorem, we can make this the free vacuum expectation value of the Wick expansion. But recall that the vacuum expectation value of any normal-ordered product is zero[11]. Therefore, any term with uncontracted fields is zero.

$$\langle 0|\mathrm{T}\big[\hat{\phi}_I(x_1)\,\hat{\phi}_I(x_2)\,\hat{\phi}_I(x_3)\cdots\big]|0\rangle = \langle 0| \begin{pmatrix}\text{all complete}\\ \text{contractions}\end{pmatrix} |0\rangle$$

[11] See equation 119

As before, for brevity, let ϕ_i denote $\hat{\phi}_I(x_i)$. Note that now we are also dropping the I to denote interaction picture field. When doing contractions we will always be dealing with interaction picture fields. For example,

$$\langle 0|\mathrm{T}\big[\phi_1\,\phi_2\,\phi_3\,\phi_4\big]|0\rangle = \langle 0|\wick{\c1\phi_1\,\c1\phi_2\,\c2\phi_3\,\c2\phi_4}|0\rangle + \langle 0|\wick{\c1\phi_1\,\c2\phi_2\,\c1\phi_3\,\c2\phi_4}|0\rangle + \langle 0|\wick{\c1\phi_1\,\c2\phi_2\,\c2\phi_3\,\c1\phi_4}|0\rangle$$

So what can we do with these completely contracted terms? From the definition of Wick contraction, equation 152, and knowing that the vacuum expectation value of any normal-ordered product is zero, we see that the vacuum expectation value of a Wick contraction is actually a Feynman propagator.

$$\begin{aligned}\langle 0|\wick{\c1\phi_1\,\c1\phi_2}|0\rangle &= \langle 0|\mathrm{T}\big[\phi_1\,\phi_2\big]|0\rangle - \cancelto{0}{\langle 0|\mathrm{N}\big[\phi_1\,\phi_2\big]|0\rangle} \\ &= i\,\Delta_F(x_1 - x_2)\end{aligned}$$

Note from equations 151 and 107, that because a Wick contraction is ultimately a commutator between creation and annihilation operators, as an operator, it is proportional to unity.

$$\wick{\c1\phi_1\,\c1\phi_2} \propto \hat{\mathbb{1}}$$

Therefore,

$$\langle 0|\wick{\c1\phi_1\,\c1\phi_2\,\c2\phi_3\,\c2\phi_4}|0\rangle = \langle 0|\wick{\c1\phi_1\,\c1\phi_2}|0\rangle\,\langle 0|\wick{\c1\phi_3\,\c1\phi_4}|0\rangle$$

and

$$\begin{aligned}\langle 0|\mathrm{T}\big[\phi_1\,\phi_2\,\phi_3\,\phi_4\big]|0\rangle = &\quad \langle 0|\wick{\c1\phi_1\,\c1\phi_2}|0\rangle\,\langle 0|\wick{\c1\phi_3\,\c1\phi_4}|0\rangle \\ &+ \langle 0|\wick{\c1\phi_1\,\c1\phi_3}|0\rangle\,\langle 0|\wick{\c1\phi_2\,\c1\phi_4}|0\rangle \\ &+ \langle 0|\wick{\c1\phi_1\,\c1\phi_4}|0\rangle\,\langle 0|\wick{\c1\phi_2\,\c1\phi_3}|0\rangle \\ = &- \Delta_F(x_1 - x_2)\,\Delta_F(x_3 - x_4) \\ &- \Delta_F(x_1 - x_3)\,\Delta_F(x_2 - x_4) \\ &- \Delta_F(x_1 - x_4)\,\Delta_F(x_2 - x_3)\end{aligned}$$

Note that the vacuum expectation value of an odd number of field operators will always be zero because it will always leave an uncontracted field.

4.3 Feynman Diagrams

So far, in this chapter, we have derived a formula for relating correlation functions to a perturbation series where each term is a free vacuum expectation value of a time-ordered product of interaction picture fields. Then we showed how these vacuum expectation values could be converted to products of Feynman propagators using Wick's theorem.

Let us go back to considering the numerator of the 2-point function, equation 148, whose first few terms we expanded previously in equation 150. Let us consider writing these terms as Wick contractions yielding Feynman propagators. The first term is already a Feynman propagator. For the next term, we need to write out the Wick expansion.

$$\begin{aligned}\langle 0|\mathrm{T}\left[\hat{\phi}_I(x_1)\,\hat{\phi}_I(x_2)\,\hat{\phi}_I^4(y_1)\right]|0\rangle &= \langle 0|\hat{\phi}_I(x_2)\,\hat{\phi}_I(x_1)\,\hat{\phi}_I(y_1)\,\hat{\phi}_I(y_1)\,\hat{\phi}_I(y_1)\,\hat{\phi}_I(y_1)|0\rangle \\ &+ \langle 0|\hat{\phi}_I(x_1)\,\hat{\phi}_I(x_2)\,\hat{\phi}_I(y_1)\,\hat{\phi}_I(y_1)\,\hat{\phi}_I(y_1)\,\hat{\phi}_I(y_1)|0\rangle \\ &+ \text{all other complete contractions} \qquad (154)\end{aligned}$$

Note that only a subset of all the complete contractions involve unique combinations of spacetime points. For example, the following contractions are equivalent.

$$\begin{aligned}&\langle 0|\hat{\phi}_I(x_1)\,\hat{\phi}_I(x_2)\,\hat{\phi}_I(y_1)\,\hat{\phi}_I(y_1)\,\hat{\phi}_I(y_1)\,\hat{\phi}_I(y_1)|0\rangle \\ =\;&\langle 0|\hat{\phi}_I(x_1)\,\hat{\phi}_I(x_2)\,\hat{\phi}_I(y_1)\,\hat{\phi}_I(y_1)\,\hat{\phi}_I(y_1)\,\hat{\phi}_I(y_1)|0\rangle\end{aligned}$$

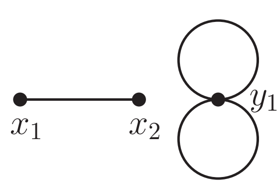
To help organize things, we will now introduce a diagrammatic notation, called **Feynman diagrams**, for keeping track of the possible unique contractions. To denote the contraction of two fields, we draw a line connecting two points representing the spacetime points at which the fields are evaluated.

$$\equiv \langle 0|\hat{\phi}_I(y)\,\hat{\phi}_I(x)|0\rangle = i\,\Delta_F(x-y)$$

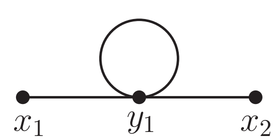
This line equivalently represents a Feynman propagator from one spacetime point to the other. Stitching together these lines, we can summarize the Wick expansions. Using this notation, the first-order term in equation 154 reads as follows.

$$\frac{1}{1!}\left(\frac{-i\,\lambda}{4!}\right)\int d^4y_1\;\langle 0|\mathrm{T}\left[\hat{\phi}_I(x_1)\,\hat{\phi}_I(x_2)\,\hat{\phi}_I^4(y_1)\right]|0\rangle$$

$$= \frac{3}{1!\,4!}$$



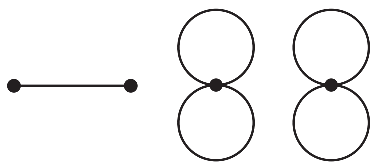
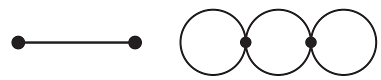
$$+\;\frac{12}{1!\,4!} \qquad (155)$$

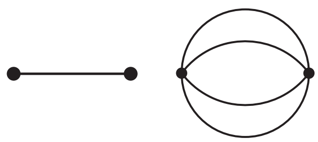
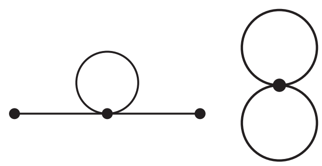


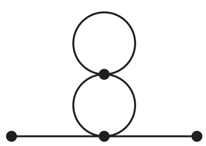
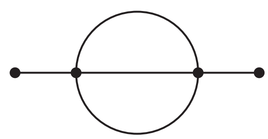
Each internal vertex has an implied factor of $(-i\,\lambda)$ and is integrated over all spacetime. For now, the rest of the numerical factors are left out front. There are only two diagrams because there are only two unique topologies for pairing the operators in contractions. The factors of 3 and 12 count the number of Wick contractions that generate the same diagram. The first term in equation 154 is an example of a contraction contributing to the first diagram, and the second term in equation 154 is an example of a contraction contributing to the second.

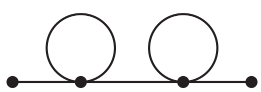
Moving on to the next (second-order) term in the numerator of the 2-point function, equation 150, we have the following diagrams.

$$\frac{1}{2!}\left(\frac{-i\,\lambda}{4!}\right)^2 \int d^4y_1\, d^4y_2\, \langle 0|\mathrm{T}\left[\hat{\phi}_I(x_1)\,\hat{\phi}_I(x_2)\,\hat{\phi}_I^4(y_1)\,\hat{\phi}_I^4(y_2)\right]|0\rangle =$$

$$\frac{9}{2!\,(4!)^2}$$  $$+\frac{72}{2!\,(4!)^2}$$ 

$$+\frac{24}{2!\,(4!)^2}$$  $$+\frac{72}{2!\,(4!)^2}$$ 

$$+\frac{288}{2!\,(4!)^2}$$  $$+\frac{192}{2!\,(4!)^2}$$ 

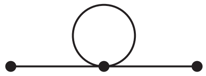
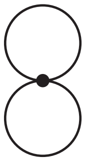
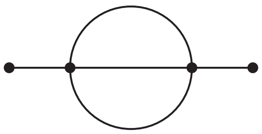
$$+\frac{288}{2!\,(4!)^2}$$  (156)

You are probably daunted by the combinatoric coefficients in front of these diagrams, but the following trick makes calculating them much more simple.

- In a diagram with n vertices, if each vertex is topologically unique, then there are $n!$ combinatoric ways for exchanging the labels of each vertex which give the same diagram. This $n!$ cancels the $\frac{1}{n!}$ in the Taylor series.

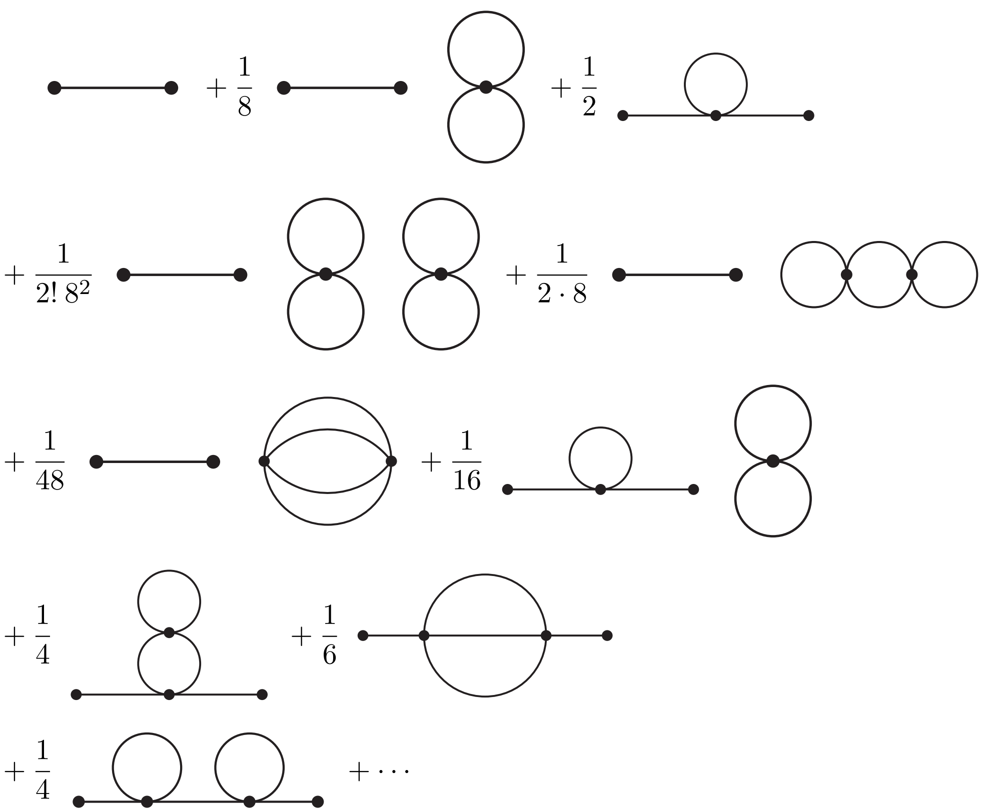
- If each line coming from a vertex connects that vertex to a unique spacetime point, then there are 4! ways of exchanging the lines which give the same diagram. This 4! cancels the $\frac{1}{4!}$ in the $\frac{\lambda}{4!}$ coupling. This was the reason for putting this factor in the coupling coefficient.

If we assume these rules but a diagram has symmetries that violate these rules, then we would have over counted the Wick contractions contributing to that diagram by some **symmetry factor**. This symmetry factor is the number of ways of interchanging components of the diagram without changing the topology of the diagram. Therefore the overall coefficient of a diagram is the reciprocal of its symmetry factor. Examples of symmetry factors are as follows.

	$S = 2$	flipping the loop
	$S = 2 \cdot 2 \cdot 2 = 8$	flipping each loop and flipping over the whole diagram
	$S = 3! = 6$	exchanging the internal lines joining the vertices

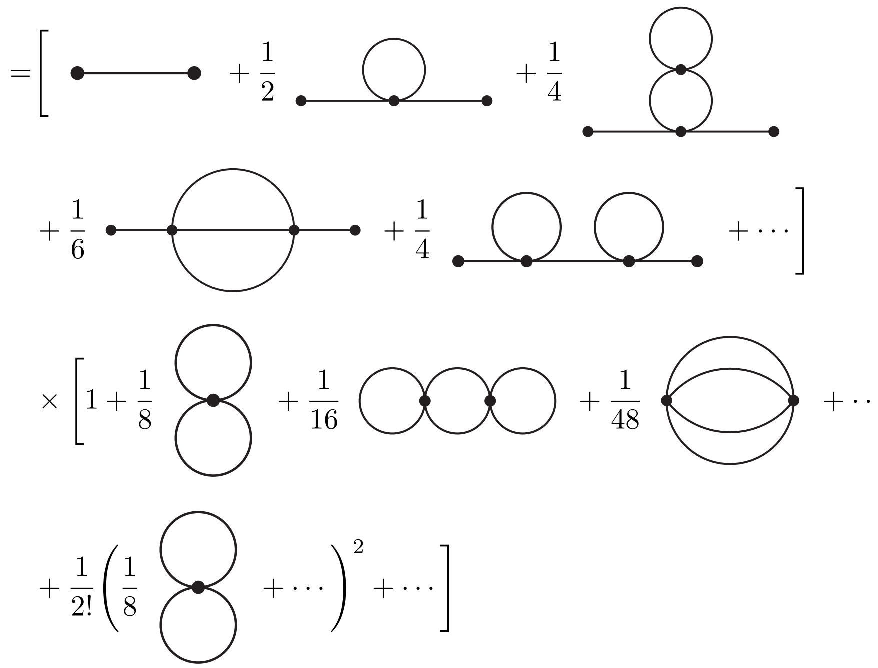
Combining all these terms up to second-order, we have the following for the numerator of the 2-point function, equation 150.

$$\langle 0|\mathrm{T}\left[\hat{\phi}_I(x_1)\,\hat{\phi}_I(x_2)\,\exp\left(\frac{-i\,\lambda}{4!}\int d^4y\,\hat{\phi}_I^4(y)\right)\right]|0\rangle =$$



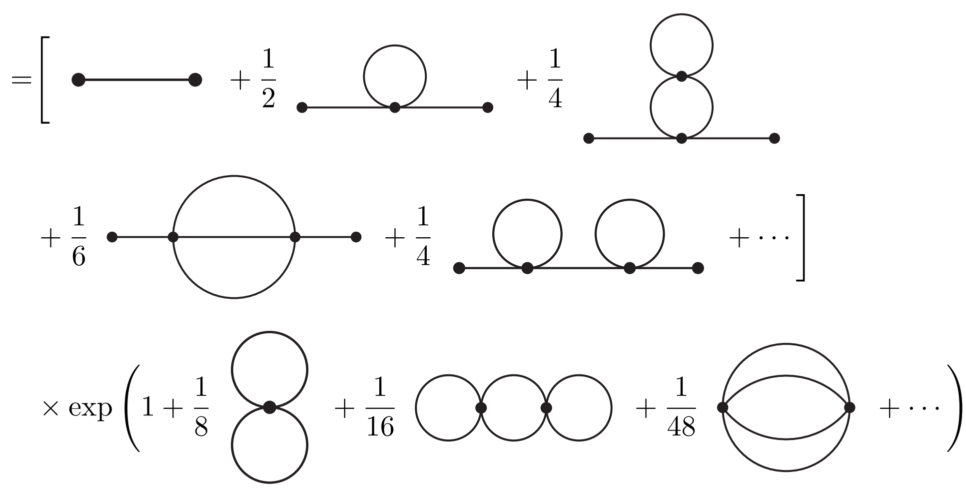
(157)

Looking at this expression carefully, notice that the parts of diagrams that are not connected to any external spacetime points, called **vacuum bubbles**, can be factored as follows.



(158)

This motivates the amazing fact that the factor of *vacuum bubbles forms an exponential.*

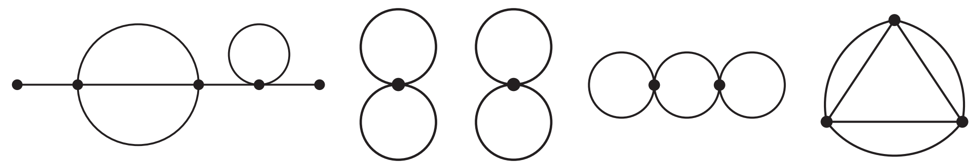


$$= \sum \left(\begin{matrix} \text{external} \\ \text{diagrams} \end{matrix} \right) \exp \left(\sum \left(\begin{matrix} \text{vacuum} \\ \text{bubbles} \end{matrix} \right) \right) \tag{159}$$

Note that the 1/2! symmetry factor from exchanging the vacuum vertices in the term with the double figure-eight bubbles in equation 158, factored out to given the proper coefficient for the exponential. Also note that after pulling out this 1/2!, the figure-eight bubble was left with the reciprocal of its own symmetry factor as a coefficient. This shows that even in the exponent, the coefficient of any diagram is just the reciprocal of its symmetry factor.

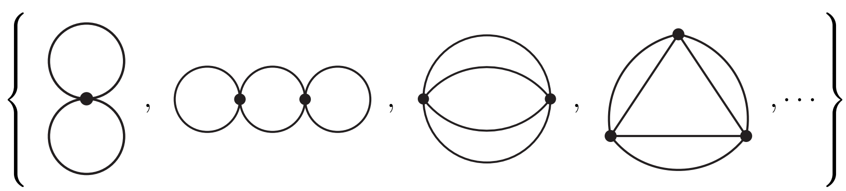
Therefore, people often write these sums as just the sum of the possible diagrams, leaving the symmetry factor implied because it can be determined directly from the diagram. *From now on, we leave the symmetry factors of individual connected diagrams implied.*

We prove that the vacuum bubbles factor into an exponential as follows. In general, a diagram has a part with external points and the product of some set of vacuum bubbles. For example



Since the set of possible vacuum bubbles is enumerable, albeit infinite, we can label the i-th vacuum bubble by V_i, including its symmetry factor.

$V_i \in$



Then the value of a diagram with a given external part of a diagram and a given set of vacuum bubbles is

$$\begin{pmatrix} \text{a given external part} \\ \text{with } \{n_i\} \text{ bubbles} \end{pmatrix} = (\text{external part}) \left(\prod_i \frac{1}{n_i!} V_i^{n_i} \right) \tag{160}$$

where $\{n_i\}$ is an ordred set of integers indicating that there are n_i coppies of the V_i bubble in the total diagram. That external part can be can be paired with every possible combination of vacuum bubbles, so we sum over every possible ordered set $\{n_i\}$.

$$\begin{pmatrix} \text{a given external part} \\ \text{with all possible bubbles} \end{pmatrix} = (\text{external part}) \sum_{\{n_i\}} \left(\prod_i \frac{1}{n_i!} V_i^{n_i} \right) \tag{161}$$

Finally, we have to sum over every possible external diagram, each term having a sum over all sets of vacuum bubbles, so we factor the vacuum bubbles out.

$$\begin{aligned} &\langle 0 | \mathrm{T} \left[\hat{\phi}_I(x_1)\, \hat{\phi}_I(x_2)\, \exp\left(\frac{-i\,\lambda}{4!} \int d^4y\, \hat{\phi}_I^4(y) \right) \right] |0\rangle \\ &= \sum \begin{pmatrix} \text{external} \\ \text{diagrams} \end{pmatrix} \sum_{\{n_i\}} \left(\prod_i \frac{1}{n_i!} V_i^{n_i} \right) \end{aligned}$$

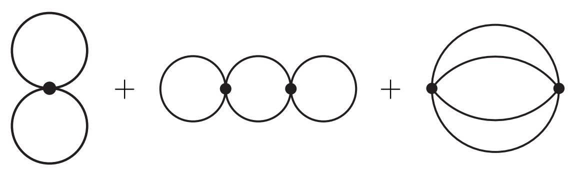
This sum over all possibles sets $\{n_i\}$ can be written as a product of sums over all values n_i, the cross terms giving every possible combination.

$$
\begin{aligned}
&= \sum \begin{pmatrix} \text{external} \\ \text{diagrams} \end{pmatrix} \left(\sum_{n_1} \frac{1}{n_1!} V_1^{n_1}\right) \left(\sum_{n_2} \frac{1}{n_2!} V_2^{n_2}\right) \left(\sum_{n_3} \frac{1}{n_3!} V_3^{n_3}\right) \cdots \\
&= \sum \begin{pmatrix} \text{external} \\ \text{diagrams} \end{pmatrix} \prod_i \left(\sum_{n_i} \frac{1}{n_i!} V_i^{n_i}\right) \\
&= \sum \begin{pmatrix} \text{external} \\ \text{diagrams} \end{pmatrix} \prod_i \exp(V_i) \\
&= \sum \begin{pmatrix} \text{external} \\ \text{diagrams} \end{pmatrix} \exp\left(\sum_i V_i\right)
\end{aligned}
$$

Therefore we have verified equation 159.

We have simplified the numerator of the 2-point function as much as possible. Now we turn our attention to the denominator. Using the same argument we used on the numerator: expanding out the exponential, writing the Wick contractions for each term, converting each Wick contraction to a Feynman diagram, and regrouping the terms into an exponential; we would find that the denominator is simply the exponential of the sum of all vacuum bubbles.

$$
\langle 0|\mathrm{T}\left[\exp\left(\frac{-i\,\lambda}{4!}\int d^4y\,\hat{\phi}_I^4(y)\right)\right]|0\rangle
$$

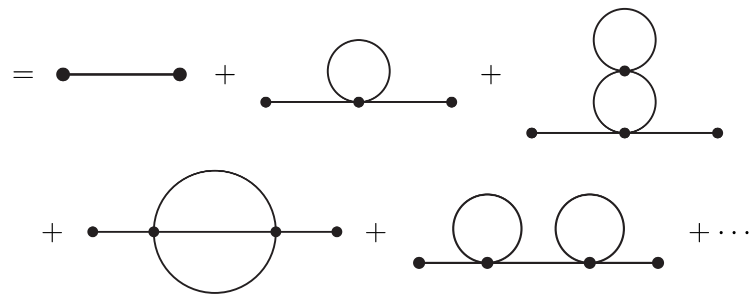
$= \exp\Big($  $+ \cdots \Big)$ (162)

Therefore[12], it *cancels* the exponential in the numerator. To summarize, for the 2-point

[12] One may be tempted to look at equation 162 and think that on the left side one can bring the time ordering and expectation value into the exponential, and conclude that the arguments of the exponentials themselves are equal, but this is not so. Recall that the left exponential is actually an abuse of notation, a mnemonic to denote the entire Dyson series, equation 141. Each successive term in the "exponential" on the left has an additional integration over a new spacetime point. Without expanding the series to reveal all the integrals, we would never get additional vertices for the higher order vacuum bubbles. The exponential on the right, however, is an exponential in the true sense. Recall that the Feynman diagrams really represent a product of Feynman propagators, which is just a scalar function. So the argument of this exponential is just a number, even though doing the integrals shows that it is actually infinite.

function we have.

$$G^{(2)}(x_1, x_2) \equiv \langle \Omega | \hat{\phi}(x_1)\, \hat{\phi}(x_2) | \Omega \rangle = \sum \begin{pmatrix} \text{external} \\ \text{diagrams} \end{pmatrix}$$

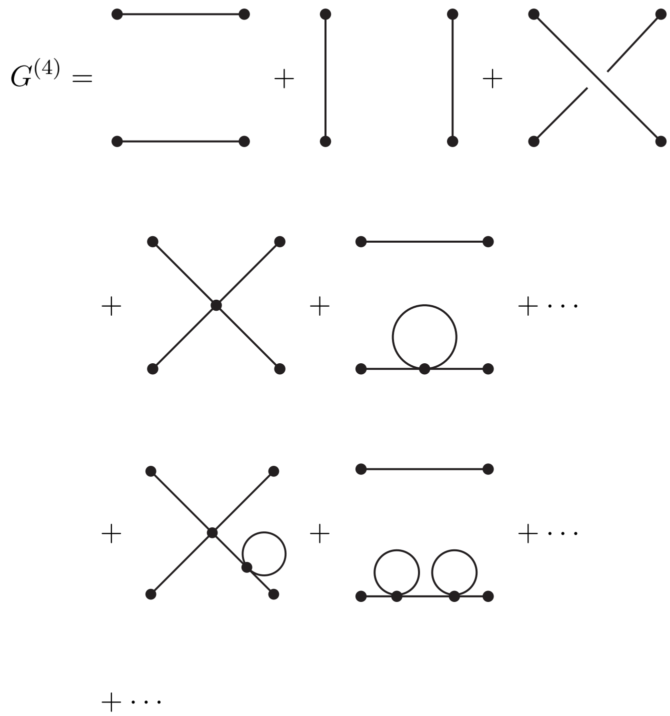


(163)

This factoring and canceling of the vacuum bubbles works the same way for all n-point functions, so in general, the n-point function is given by

$$G^{(n)}(x_1, x_2, \ldots, x_n) \equiv \langle \Omega | \hat{\phi}(x_1)\hat{\phi}(x_2) \cdots \hat{\phi}(x_n) | \Omega \rangle = \sum \begin{pmatrix} \text{diagrams with} \\ n \text{ external points} \end{pmatrix} \tag{164}$$

This sum is over all possible diagrams where every connected part of the diagram has an external point, but these connected parts can be disconnected from one another. Examples of terms like this are in the 4-point function, whose series is as follows.



Acknowledgments

I gratefully acknowledge the help and perspective of many others in preparing this document. All the remaining errors and defficiencies in clarity are my fault.

Burt Ovrut, Peskin and Schroeder (1995), David Tong.

References