Physics 615, Lectures 1-3 Sept. 5–12, 2013

Introduction



# 1 Introduction

According to the catalogue, this course is "Overview of Quantum Field Theory". Why "Overview"? When we teach Classical Mechanics or Quantum Mechanics, we teach the groundwork of those subjects, at least up to some point, and thus we call the courses "Introductions to" rather than "Overviews of". Giving an overview is an unusual way to start teaching one of the broad conceptual frameworks to Physicists. We know that these theoretical frameworks need to built on a firm conceptual basis. All physicists need that understanding and so we don't try to include a preview of time dependant perturbation theory in freshman physics courses.

Quantum Field theory is also a broad conceptual framework underlying much of how physicists understand their field. It is, except possibly for the much more abstract string theory, the only framework for understanding elementary particle physics at energies high enough for relativity to matter. In ordinary Quantum Mechanics, the basic operators are the degrees of freedom, usually the positions and momenta of each particle. With sufficient energy available, particles can be created, so the coordinates describing the positions or spins of a list of individual particles cannot suffice to describe the state of a system, and thus the quantum mechanical framework used in nonrelativistic physics is inadequate.

Quantum field theory is also extensively used in condensed matter physics, even though here there is not enough energy for electrons and positrons to be created. In condensed matter physics one often describes the system not in terms of fundamental particles or even nuclei or atoms, but in terms of effective particles such as holes, phonons or excitons, and these are not conserved.

Even in atomic physics, because intermediate states in perturbation theory may not conserve energy, the possibility of particle creation has effects, such as the Lamb shift in hydrogenic atoms, and larger effects in the inner shells of heavy atoms, for which the potential energies may be comparable to the rest masses of the electrons.

So, if field theory is the language of so much of modern physics, why an overview rather than a careful solid foundation-laying introduction? The motivation is that right from the beginning, QFT raises a host of issues which are quite complex. Unlike in classical mechanics or nonrelativistic quantum mechanics, there are no nontrivial problems which can be solved exactly. For example, even though we can write a "complete" theory of electrodynamics, complete in the sense that it would be if there were not other interactions, which is quite adequate for understanding a great deal of observed physics, we still cannot do any exact calculations with that theory, except if we assume there are no interactions at all. What we can do, very successfully, are calculations in perturbation theory, treating the interaction parameter, which is the charge of the electron, as a small parameter in terms of which we can calculate scattering amplitudes as power series expansions. But these calculations are both complicated and raise some very difficult issues, with which you should be acquainted, even though you may not be prepared to study these ideas in detail. Some examples:

- Renormalization is required to make sense of all but the lowest level approximation to scattering amplitudes. The renormalization **group** is important in understanding critical phenomena and also for the running coupling constants one must use even in low order calculations of high energy scattering experiments.

- Gauged symmetry groups are responsible for the standard model of high energy physics and also provide models in CM, but a full discussion of their quantization is very involved.

- Spontaneous symmetry breaking plays a crucial role in the standard model, as well as in condensed matter (*e.g.* ferromagnetism), and is probably responsible for the inflation that set up the initial conditions for cosmology.

- The Higgs model is an elaboration on spontaneous symmetry breaking in theories with gauged symmetry groups. Higgs models are essential in both high energy (the standard model) and condensed matter physics.

The text, Peskin and Schroeder's "Introduction to Quantum Field Theory", does indeed give a good **introduction** laying out these ideas, but it is far too much to teach in one semester. So this course will be a compromise, a fairly solid foundation to the beginning of quantum field theory, together with a more qualitative or sketchy overview of the more advanced ideas that I have just mentioned. We will learn how to use Feynman graphs to describe scattering amplitudes in perturbation theory, at least at the level that experimentalists need for most purposes.

For today, we will have a brief review of Classical Mechanics in which we make sure we learn to deal with fields. Then we will discuss the limitations of particle quantum mechanics and begin to develop the quantum mechanics of fields. This process has been called **second quantization**. In the first quantization, in the 1920's, we learned to replace the phase space variables, *i.e.* the positions and momenta of individual particles which describe the classical mechanics of the system, by operators operating on the wave function, an ordinary function of the positions. In Quantum mechanics it is the wave function, rather than the coordinates, that describes the state of the system. But in quantum field theory this wavefunction becomes a field which is then quantized, and we will treat the degrees of freedom of this field as operators.

I think this name, second quantization, is losing its appeal, and in a sense it is misleading anyway. After all, historically the first thing to be quantized was the electromagnetic field inside a black box. Here we were used to the idea that light was described by fields, which could be considered classical, while it was the particle nature of photons that was considered quantum mechanical. For electrons, the particle nature was classical and the wave function, a field $\psi(\vec{x}, t)$ was considered quantum mechanical. But the de Broglie relationship between momentum and wavelength is the same for photons and electrons, and the Schrödinger Equation can be derived by applying de Broglie to the Newtonian equations just as Maxwell's equations relates to electromagnetic waves.

Even in classical mechanics, relativity does not get along very well with a particle description. There is no problem for free point particles, or for describing collisions at a point. But for interesting particle mechanics we need potential energies which are functions of the separation of the particles $\vec{x}_1 - \vec{x}_2$, and not just delta functions $\delta^3(\vec{x}_1 - \vec{x}_2)$. But such a potential describes action at a distance, the force on particle 1 at time $t$ depending on the position of particle 2 **at the same time** $t$, which is impossible for particle 1 to know in relativity. The only relativistic interactions between separated particles known to classical mechanics are electromagnetic (and, in a sense, general relativistic gravity), and these interactions cannot be understood in general without the introduction of fields.

[I am not going to cover Chapter 1 of the book — read it if you like. I will be covering, in the first four lectures, the material of Chapter 2, though somewhat differently.]

# 2 Review of Classical Mechanics

Let us very briefly review the fundamentals of mechanics. In describing a physical situation, the first thing one must do is decide what the coordinates are. In classical mechanics of a discrete set of degrees of freedom, we will have coordinates $q_j(t)$ which can depend on time. For the systems we wish to deal with, the physics can be described by a Lagrangian

$$L(t) = L(\{q_j(t)\}, \{\dot{q}_j(t)\}, t)$$

which is a function of the coordinates and their first time derivatives at a moment of time. The dynamics of the system tells how $q_j$ develops in time. That is, it gives the path $\Gamma : t \in [t_1, t_2] \mapsto q_j(t)$ that the physical system will actually take through coordinate space as a function of time. This is determined by a functional called the **action**

$$S_\Gamma = \int_{t_1}^{t_2} L(\{q_j(t)\}, \{\dot{q}_j(t)\}, t)\, dt,$$

where the integral is to be done with $q$ and $\dot{q}$ evaluated along an arbitrary path $\Gamma$. Hamilton's principle tells us the actual physical path will be one which is stationary under small variations in the path, keeping the endpoints $q_j(t_1)$ and $q_j(t_2)$ fixed. Considering an arbitrary infinitesimal variation of the path, $\delta q_j(t)$, (with $\delta q_j(t_1) = \delta q_j(t_2) = 0$), and asking that $\delta S = 0$, gives the Lagrange equations,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0.$$

We should note here that the space in which q lives is not necessarily ordinary three dimensional Euclidean space, or even a Euclidean space of any number of dimensions, though for single particle motion it is ordinary space. For example, for a pendulum swinging about a fixed axis, the sole coordinate may be $\theta$, the angle with respect to the downward direction, which lives on a circle (with $\theta = 0$ and $\theta = 2\pi$ the same point). If it is swinging from a fixed point rather than a fixed axis, its coordinates are the polar angle $\theta$ and the azimuthal angle $\phi$, which together live on a two-sphere $S^2$, that is, the surface of a three dimensional ball.

Classical mechanics of a conservative system can also be described in terms of a Hamiltonian, which depends on momenta $p_j$ as well as the coordinates, but does not depend on the time derivatives of the coordinates:

$$H(\{q_j(t)\}, \{p_j(t)\}, t).$$

The **canonical momentum** $p_j$ conjugate to $q_j$ is given in the Lagrangian formulation by

$$p_j = \frac{\partial L}{\partial \dot{q}_j}. \tag{1}$$

Eq. (1) is known as the **constitutive equation**. Note that these canonical momenta, one for each degree of freedom, are to be distinguished from the 3 (or 4) components of the total momentum, except for a single particle problem described by cartesian coordinates. Usually this expression for the canonical momenta $p_i$ in terms of $\{q_j\}$ and $\{\dot{q}_j\}$ can be inverted to give $\dot{q}_j(t)$ in terms of $\{q_k(t)\}$ and $\{p_k(t)\}$, in which case one can define the Hamiltonian from the Lagrangian by

$$H(\{q_j(t)\}, \{p_j(t)\}, t) = \sum_k \dot{q}_k(t) p_k(t) - L(\{q_j(t)\}, \{\dot{q}_j(t)\}, t),$$

where it is understood that on the right hand side the $\dot{q}_j$'s are to be replaced by their values in terms of $q$ and $p$. When Eq. (1) is not solvable for $\dot{q}$, we have a complicated situation related to gauge invariance. We will see that this situation does arise for the electromagnetic field.

The equations of motion in Hamiltonian form are

$$\dot{q}_j = \frac{\partial H}{\partial p_j}, \qquad -\dot{p}_j = \frac{\partial H}{\partial q_j},$$

which determine the classical path the system takes through phase space, $(q(t), p(t))$.

We can also relate any explicit time dependences of $L$ and $H$ by

$$-\frac{\partial L}{\partial t} = \frac{\partial H}{\partial t}.$$

The Hamiltonian is a function on extended phase space, $\{q_j, p_j, t\}$ while the Lagrangian is not quite a function on coordinate space, because it depends of $\dot{q}_j$ as well as $q_j$. The Hamiltonian equations of motion can be derived from a variational principle in which one varies a path $\Gamma$ through phase space $\Gamma : t \mapsto (q(t), p(t))$, varying $q(t)$ and $p(t)$ independently to extremize the action

$$S = \int \left( \sum_j p_j \dot{q}_j - H(q, p, t) \right) dt,$$

where $\dot{q}$ is evaluated from the path $q(t)$ and is unrelated to $p(t)$.

At the beginning of a course on classical or quantum mechanics we deal with a system of discrete degrees of freedom, $q_j$. For electromagnetism, however, the degrees of freedom are not only the positions of the discrete charged particles, but also the electric and magnetic **fields** at each point in space, as a function of time. As we shall argue shortly, even for ordinary particles we are going to need to deal with fields, where there is one (or perhaps several) degree(s) of freedom for each point in space, at each time.

To make sure we understand formally how to deal with fields, let's begin with the simple example of the continuum limit of a one dimensional lattice of masses on springs. Each lattice point $x_j = j\,a$ (for integer $j$, that is $j \in \mathbb{Z}$) has a mass $m$ which may be displaced from its equilibrium position $x_j$ by an amount $\eta_j(t)$. Between neighboring points there are springs of equilibrium length $a$ and spring constant $k$, so the kinetic and potential energies of the system are

$$T = \frac{1}{2} m \sum_j \dot{\eta}_j^2, \qquad V(\{\eta_j\}) = \frac{1}{2} k \sum_j \left(\eta_{j+1} - \eta_j\right)^2 .$$

In the continuum limit, it is more appropriate to use the linear mass density $\mu = m/a$, Young's modulus $Y$ and the Lagrangian density. Young's modulus is the force required to produce a unit extension per unit length, and as the force required to stretch one spring, of length $a$, is $F = k(\eta_{j+1} - \eta_j)$, we have $F = Y(\eta_{j+1} - \eta_j)/a$, so $k = Y/a$. Thus we may write the lagrangian as

$$L = a \sum_j L_j = a \left\{ \frac{1}{2} \mu \sum_j \dot{\eta}_j^2 - \frac{1}{2} Y \left( \frac{\eta_{j+1} - \eta_j}{a} \right)^2 \right\}.$$

The Lagrangian density in the continuum limit is

$$\mathcal{L}(x) = \lim_{a \to 0} L_{x/a} = \frac{1}{2} \mu \dot{\eta}^2 - \frac{1}{2} Y \left( \frac{\partial \eta}{\partial x} \right)^2 .$$

and $L = \int \mathcal{L}(x)dx$.

This Lagrangian, however, will not be of much use until we figure out what is meant by varying it with respect to each dynamical degree of freedom or its corresponding velocity. In the discrete case we have the canonical momenta $p_j = \partial L/\partial\dot\eta_j$, where the derivative requires holding all $\dot\eta_j$ fixed, for $j \neq i$, as well as all $\eta_k$ fixed. This extracts one term from the sum $\frac{1}{2}\mu\sum a\dot\eta_j^2$, and this would appear to vanish in the limit $a \to 0$. Instead, we define the canonical momentum as a density, $p_j \to ap(x = ja)$, so

$$p(x = ja) = \lim_{a\to 0}\frac{1}{a}\frac{\partial}{\partial\dot\eta_j}\sum_j a\,\mathcal{L}(\eta(x),\dot\eta(x),x)\big|_{x=aj}.$$

We may think of the last part of this limit,

$$\lim_{a\to 0}\sum_j a\,\mathcal{L}(\eta(x),\dot\eta(x),x)\big|_{x=aj} = \int dx\,\mathcal{L}(\eta(x),\dot\eta(x),x),$$

if we also define a limiting operation

$$\lim_{a\to 0}\frac{1}{a}\frac{\partial}{\partial\dot\eta_j} \to \frac{\delta}{\delta\dot\eta(x)},$$

and similarly for $\frac{1}{a}\frac{\partial}{\partial\eta_j}$, which act on functionals of $\eta(x)$ and $\dot\eta(x)$ by

$$\frac{\delta\eta(x_1)}{\delta\eta(x_2)} = \delta(x_1 - x_2), \quad \frac{\delta\dot\eta(x_1)}{\delta\eta(x_2)} = \frac{\delta\eta(x_1)}{\delta\dot\eta(x_2)} = 0, \quad \frac{\delta\dot\eta(x_1)}{\delta\dot\eta(x_2)} = \delta(x_1 - x_2).$$

Here $\delta(x' - x)$ is the **Dirac delta function**, defined by its integral,

$$\int_{x_1}^{x_2} f(x')\,\delta(x' - x)\,dx' = f(x)$$

for any function $f(x)$, provided $x$ is contained in the interval $(x_1, x_2)$. Thus

$$p(x) = \frac{\delta}{\delta\dot\eta(x)}\int_0^\ell dx'\frac{1}{2}\mu\dot\eta^2(x',t) = \int_0^\ell dx'\mu\dot\eta(x',t)\delta(x' - x) = \mu\dot\eta(x,t).$$

We also need to evaluate

$$\frac{\delta}{\delta\eta(x)}L = \frac{\delta}{\delta\eta(x)}\int_0^\ell dx'\frac{-Y}{2}\left(\frac{\partial\eta}{\partial x}\right)^2_{x=x'}.$$

For this we need

$$\frac{\delta}{\delta\eta(x)}\frac{\partial\eta(x')}{\partial x'} = \frac{\partial}{\partial x'}\delta(x'-x) := \delta'(x'-x).$$

The generalized function $\delta'(x'-x)$ is also defined by its integral,

$$\begin{aligned}
\int_{x_1}^{x_2} f(x')\delta'(x'-x)dx' &= \int_{x_1}^{x_2} f(x')\frac{\partial}{\partial x'}\delta(x'-x)dx' \\
&= f(x')\delta(x'-x)\big|_{x_1}^{x_2} - \int_{x_1}^{x_2} dx'\,\frac{\partial f}{\partial x'}\delta(x'-x) \\
&= -\frac{\partial f}{\partial x}(x),
\end{aligned}$$

where after integration by parts the surface term is dropped because $\delta(x-x')=0$ for $x \neq x'$, which it is for $x' = x_1, x_2$ if $x \in (x_1, x_2)$. Thus

$$\frac{\delta}{\delta\eta(x)}L = -\int_0^\ell dx' Y\frac{\partial\eta}{\partial x}(x')\delta'(x'-x) = Y\frac{\partial^2\eta}{\partial x^2},$$

and Lagrange's equations give the wave equation

$$\mu\ddot{\eta}(x,t) - Y\frac{\partial^2\eta}{\partial x^2} = 0.$$

## 2.1 Higher dimensions

Now let us generalize to more than one spatial dimension, say $D-1$, reserving $D$ to represent the dimension of space-time. Of course we are usually interested in $D=4$, to describe the apparent real world, but other dimensions will come up in condensed matter, in dimensional regularization, and in string theory. We will use relativistic notation, with coordinates $x^\mu$, with $\mu = 0, 1, \ldots D-1$, and with $x^0 = ct = t$, because we use units where the speed of light is 1. Thus times and distances are both measured in meters. We are distinguishing contravariant vectors $x^\mu$ from covariant vectors $x_\mu$, related by a metric tensor

$$g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad \text{with } x^\mu = g^{\mu\nu}x_\nu,$$

and $x_\mu = g_{\mu\nu}x^\nu$. We are also using the "Einstein summation convention", that an index repeated up and down is automatically summed. The momenta are naturally covariant, for in quantum mechanics

$$p_\mu \sim i\hbar \frac{\partial}{\partial x^\mu},$$

though if we write it in nonrelativistic notation,

$$p^\mu = (E, \vec{p}) \sim \left(i\partial/\partial t, -i\vec{\nabla}\right).$$

Note

$$\frac{\partial}{\partial x^\mu} x^\nu = \delta^\nu_\mu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The spatial part of a contravariant vector will be represented by $\vec{x}$ (boldface in the book, whose notation we otherwise follow). The dot product of two $D$-vectors is understood to use the Minkowski metric, $p^2 := g^{\mu\nu}p_\mu p_\nu = p_0^2 - \vec{p}^2 = E^2 - \vec{p}^2$, which for a particle is $m^2$, classically at least.

Let us consider a scalar field $\phi(\vec{x}, t)$. As for the chain of masses on springs, the Lagrangian will be a spatial integral of the Lagrangian density,

$$L(t) = \int \mathcal{L}\, d^{D-1}x.$$

As for the chain, the Lagrangian density will depend on spatial as well as temporal first derivatives of $\phi$, and may also depend on $\phi$ itself:

$$\mathcal{L} = \mathcal{L}\left(\phi, \partial_\mu\phi, x^\mu\right),$$

where we are switching to relativistic notation, with

$$\partial_\mu \phi = \frac{\partial}{\partial x^\mu}\phi = \left(\dot{\phi}, \vec{\nabla}\phi\right).$$

The canonical momentum density we will[1] call $\pi(\vec{x}, t)$:

$$\pi(\vec{x}, t) := \frac{\delta L}{\delta \dot{\phi}(\vec{x}, t)},$$

[1]Peskin defines $\pi(\vec{x}, t) := \partial\mathcal{L}/\partial\dot{\phi}(\vec{x})$ which I agree with, but with some explanation.

and the hamiltonian is

$$\begin{aligned} H &= \int \mathcal{H}(\vec{x},t)\, d^{D-1}x \\ &= \int d^{D-1}x \left[\pi(\vec{x},t)\dot{\phi}(\vec{x},t) - \mathcal{L}\left(\phi(\vec{x},t), \dot{\phi}(\vec{x},t), \nabla\phi(\vec{x},t), \vec{x}, t\right)\right]. \end{aligned}$$

Before we give an example, let us note that we have not taken a very relativistic attitude towards our Lagrangian density, for the momentum density $\pi(\vec{x},t)$ is its variation with respect to the time derivative of $\phi$, but we haven't varied with respect to the spatial derivatives. We might consider the canonical momentum $\pi(\vec{x})$ as one component of a four vector

$$\frac{\delta}{\delta(\partial_\mu\phi(\vec{x},t))}L(t) = \int d^{D-1}x' \frac{\delta\mathcal{L}(\vec{x}\,',t)}{\delta(\partial_\mu\phi(\vec{x},t))}.$$

Our procedures do not fully treat $\delta/\delta(\partial_\mu\phi(\vec{x},t))$ covariantly, however, for we have been treating $\phi$ and $\dot{\phi}$ as completely independant variables, while $\phi$ and $\nabla\phi$ are not. Thus the $\delta$ variation above is not well defined.

As for the mass chain, we have

$$\frac{\delta}{\delta\phi(\vec{x},t)}\frac{\partial\phi(\vec{x}\,',t)}{\partial x'^j} = \frac{\partial}{\partial x'^j}\delta^{D-1}(x'-x),$$

where $\delta^{D-1}(\vec{x}\,'-\vec{x})$ is the $D-1$ dimensional version of the Dirac delta, zero unless $\vec{x}\,'=\vec{x}$ with $\int d^{D-1}x' f(\vec{x}\,')\delta^{D-1}(\vec{x}\,'-\vec{x}) = f(\vec{x})$. So we are not treating $\partial_j\phi$ as independent of $\phi$.

If we considered a different form of variation, $\bar{\delta}$, in which $\phi$ and $\vec{\nabla}\phi$ are considered independent and the Lagrangian density is $\mathcal{L} = \mathcal{L}(\phi,\dot{\phi},\vec{\nabla}\phi)$, then our original $\delta$ variation can be expressed in terms of $\bar{\delta}$, with

$$\begin{aligned} \frac{\delta L}{\delta\phi(\vec{x})} &= \frac{\delta}{\delta\phi(\vec{x})}\int d^{D-1}x' \mathcal{L}(\phi,\dot{\phi},\vec{\nabla}\phi)(\vec{x}\,') \\ &= \int d^{D-1}x' \frac{\delta}{\delta\phi(\vec{x})}\mathcal{L}(\phi,\dot{\phi},\vec{\nabla}\phi)(\vec{x}\,') \\ &= \int d^{D-1}x' \Bigg\{ \frac{\bar{\delta}}{\bar{\delta}\phi(\vec{x})}\mathcal{L}(\phi,\dot{\phi},\vec{\nabla}\phi)(\vec{x}\,') + \\ &\qquad \left[\frac{\bar{\delta}}{\bar{\delta}\partial_j\phi(\vec{x}\,')}\mathcal{L}(\phi,\dot{\phi},\vec{\nabla}\phi)(\vec{x}\,')\right]\frac{\delta}{\delta\phi(\vec{x})}\partial_j\phi(\vec{x}\,')\Bigg\} \end{aligned}$$

$$\begin{aligned}
&= \int d^{D-1}x' \left\{ \frac{\partial \mathcal{L}(\phi, \dot{\phi}, \vec{\nabla}\phi)}{\partial \phi}(\vec{x}\,')\delta^{D-1}(\vec{x}\,' - \vec{x}) + \right. \\
&\qquad \left. \left[\frac{\partial \mathcal{L}(\phi, \dot{\phi}, \vec{\nabla}\phi)}{\partial(\partial_j \phi)}(\vec{x}\,')\right] \frac{\partial}{\partial x'^j}\delta^{D-1}(\vec{x}\,' - \vec{x}), \right\} \\
&= \frac{\partial \mathcal{L}(\phi, \dot{\phi}, \vec{\nabla}\phi)}{\partial \phi}(\vec{x}) - \frac{\partial}{\partial x^j}\left[\frac{\partial \mathcal{L}(\phi, \dot{\phi}, \vec{\nabla}\phi)}{\partial(\partial_j \phi)}(\vec{x})\right]
\end{aligned}$$

Let us define

$$\pi^\mu(\vec{x}, t) = \frac{\bar{\delta}}{\bar{\delta}(\partial_\mu \phi(\vec{x}, t))} L(t) = \int d^{D-1}x' \frac{\bar{\delta}\mathcal{L}(\vec{x}\,', t)}{\bar{\delta}(\partial_\mu \phi(\vec{x}, t))} = \frac{\partial \mathcal{L}(\phi, \partial_\mu \phi)}{\partial(\partial_\mu \phi)}(\vec{x}, t).$$

Thus we have

$$\begin{aligned}
\frac{\delta L}{\delta \phi(\vec{x})} &= \frac{\partial \mathcal{L}(\phi, \dot{\phi}, \vec{\nabla}\phi)}{\partial \phi}(\vec{x}) - \frac{\partial}{\partial x^j}\left[\frac{\partial \mathcal{L}(\phi, \dot{\phi}, \vec{\nabla}\phi)}{\partial(\partial_j \phi)}(\vec{x})\right] \\
&= \frac{\partial \mathcal{L}}{\partial \phi(\vec{x})} - \frac{\partial}{\partial x^j}\pi^j(\vec{x})
\end{aligned}$$

Thus in writing out the Lagrange equations,

$$\begin{aligned}
\frac{d}{dt}\left(\frac{\delta L}{\delta \dot{\phi}(x)}\right) - \frac{\delta L}{\delta \phi(x)} &= \frac{d}{dt}\pi^0(\vec{x}) - \frac{\partial \mathcal{L}}{\partial \phi(x)} + \frac{\partial}{\partial x^j}\pi^j(\vec{x}) \\
&= \frac{\partial}{\partial x^\mu}\pi^\mu(\vec{x}) - \frac{\partial \mathcal{L}}{\partial \phi(x)} = 0.
\end{aligned}$$

Thus we see that, despite a rather non-covariant formulation of the laws of Lagrangian mechanics, the results are covariant equations of motion.

The covariant equations of motion are

$$\frac{\partial}{\partial x^\mu}\frac{\partial \mathcal{L}(\phi, \partial_\nu \phi)}{\partial(\partial_\mu \phi)} - \frac{\partial \mathcal{L}}{\partial \phi} = 0.$$

Our first important example will be the Klein-Gordon field, in three spatial dimensions, with a lagrangian density

$$\mathcal{L} = \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}(\nabla\phi)^2 - \frac{1}{2}m^2\phi^2 = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2}m^2\phi^2$$

This is the generalization of the mass chain to three dimensions, with an added term that depends on $\phi$ itself, if the parameter $m \neq 0$. The choice of setting the coefficient of $\dot{\phi}^2$ to $\frac{1}{2}$ may be viewed as setting the scale of $\phi$, while the coefficient of $(\nabla\phi)^2$ is then fixed if we are to have a relativistic theory. Now

$$\pi := \frac{\delta\mathcal{L}}{\delta\dot{\phi}} = \dot{\phi},$$

and more generally

$$\pi^\mu := \frac{\bar{\delta}\mathcal{L}}{\bar{\delta}\partial_\mu\phi} = \partial^\mu\phi,$$

the equations of motion are

$$\partial_\mu\partial^\mu\phi + m^2\phi = 0.$$

and the Hamiltionian density is

$$\mathcal{H} = \pi\dot{\phi} - \mathcal{L} = \frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}m^2\phi^2$$

We have several more things to discuss about classical field theory:

- Symmetries and conserved currents and charges (Noether's theorem).
- Maxwell's equations and electromagnetism, including $\mathcal{L}$ and $\mathcal{H}$, and four dimensional $F_{\mu\nu}$.

This will be postponed until the end of this chapter. First let's take a step into Quantum Mechanics.

# 3 Quantum Mechanics

In quantum mechanics the question of what path a system takes through coordinate space is no longer meaningful, nor can a system be at a point in phase space, for the Heisenburg uncertainty principle says the most localized a state can be still requires a given volume in phase space. The possible states of a system can be described in terms of wave functions on coordinate space, and quantum mechanics (Schrödinger's equation, for example) tells us how those functions evolve with time. Thus we have a unitary operator, the time-evolution operator, which gives the amplitude for a state $|\psi_1\rangle$ at one time to be in the state $|\psi_2\rangle$ a time $t$ later, $\langle\psi_2|\, e^{-iHt/\hbar}\,|\psi_1\rangle$, where we have assumed the Hamiltonian is time-independent. In fact, this transition amplitude can be understood in a functional integral formulation of quantum mechanics as a sum over all possible paths $\Gamma$ of $e^{iS_\Gamma/\hbar}$, for the system to go from one configuration to another at a later time, rather than, as in classical mechanics, choosing the one path that extremizes the action.

Actually, as we are going to be dealing only with quantum systems, it is convenient to use units with $\hbar = 1$, thereby measuring energies and (ordinary) momenta in units of inverse meters, as we have already set $c = 1$.

Let us look at a simple transition amplitude for a free nonrelativistic particle, to move from a definite position $\vec{x}_0$ at time 0 to a point $\vec{x}$ at time $t$. Here $\mathbf{H} = \vec{\mathbf{p}}^{\,2}/2m$, so the amplitude is

$$\begin{aligned}
\langle\vec{x}|\, e^{-i\mathbf{H}t}\,|\vec{x}_0\rangle &= \langle\vec{x}|\, e^{-i\mathbf{p}^2t/2m}\,|\vec{x}_0\rangle & (2)\\
&= \int \frac{d^3p}{(2\pi)^3}\,\langle\vec{x}|\, e^{-i\mathbf{p}^2t/2m}\,|\vec{p}\rangle\,\langle\vec{p}\,|\vec{x}_0\rangle & (3)\\
&= \int \frac{d^3p}{(2\pi)^3} e^{-i\vec{p}^{\,2}t/2m+i\vec{p}\cdot(\vec{x}-\vec{x}_0)} & (4)\\
&= \left(\frac{m}{2\pi it}\right)^{3/2} e^{im(\vec{x}-\vec{x}_0)^2/2t}. & (5)
\end{aligned}$$

Here[2] we are using states normalized so that $\langle\vec{x}\,'|\,\vec{x}\rangle = \delta^3(\vec{x}\,'-\vec{x})$, with momentum eigenstates normalized so that $\langle\vec{p}\,'|\,\vec{p}\rangle = (2\pi)^3\delta^3(\vec{p}\,'-\vec{p})$, and completeness is

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

and $\langle\vec{x}|\,\vec{p}\rangle = e^{ipx}$. This explains the second and third lines above. To get to the last line, it is useful to know the Gaussian integral[3]

$$\int d^Dx e^{-\vec{x}\cdot\mathbf{A}\cdot\vec{x}+\vec{B}\cdot\vec{x}} = \sqrt{\frac{\pi^D}{\det\mathbf{A}}}\, e^{\frac{1}{4}\vec{B}\cdot\mathbf{A}^{-1}\cdot\vec{B}},$$

[2]Later we will change that normalization to be consistent with Peskin and Schroeder,

$$|\vec{p}\rangle_{\rm PS} = \sqrt{2\sqrt{\vec{p}^{\,2}+m^2}}\,|\vec{p}\rangle_{\rm here}\,.$$

[3]See `http://en.wikipedia.org/wiki/Gaussian_integral`.

if $\mathbf{A}$ is a diagonalizable matrix with all eigenvalues having a positive real part. In our case we are at the boundary of this condition, but the expression still holds.

Note there is a nonzero probability that the particle will be found at any distance $|\vec{x}-\vec{x}_0|$ even after an arbitarily short time $t$. This is not surprising, for an eigenstate of position $\vec{x}_0$ contains arbitrarily high momenta, and for nonrelativistic mechanics this means arbitrarily high velocities. But what might be surprising is that the same thing happens even if we use the relativistic form for the energy, $E=\sqrt{p^2+m^2}$.

$$\begin{aligned}
\langle\vec{x}|\, e^{-i\mathbf{H}t}\, |\vec{x}_0\rangle &= \langle\vec{x}|\, e^{-i\sqrt{\mathbf{p}^2+m^2}\,t}\, |\vec{x}_0\rangle & (6)\\
&= \int \frac{d^3p}{(2\pi)^3}\, \langle\vec{x}|\, e^{-i\sqrt{\mathbf{p}^2+m^2}\,t}\, |\vec{p}\rangle\, \langle\vec{p}\,|\vec{x}_0\rangle & (7)\\
&= \frac{1}{(2\pi)^3}\int_0^\infty p^2\, dp\, e^{-i\sqrt{p^2+m^2}\,t} & (8)\\
&\qquad \times \int_0^\pi \sin\theta\, d\theta\, e^{i|p||\vec{x}-\vec{x}_0|\cos\theta}\int_0^{2\pi} d\phi & \\
&= \frac{1}{(2\pi)^2}\int_0^\infty dp\, p^2 e^{-i\sqrt{p^2+m^2}\,t}\int_0^\pi \sin\theta d\theta e^{i|p||\vec{x}-\vec{x}_0|\cos\theta} & (9)\\
&= \frac{1}{(2\pi)^2}\int_0^\infty dp\, p^2 e^{-i\sqrt{p^2+m^2}\,t}\int_{-1}^1 du\, e^{i|p||\vec{x}-\vec{x}_0|u} & (10)\\
&= \frac{1}{2\pi^2|\vec{x}-\vec{x}_0|}\int_0^\infty dp\, p\, e^{-i\sqrt{p^2+m^2}\,t}\sin(p|\vec{x}-\vec{x}_0|) & (11)
\end{aligned}$$

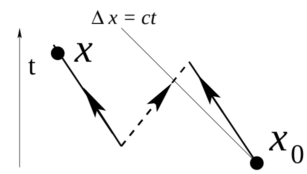
In terms of the dimensionless variables $\tau=mt$ and $r=m|\vec{x}-\vec{x}_0|$, and a dimensionless integration variable $v=p/m$, we have

$$\begin{aligned}
\langle\vec{x}|\, e^{-i\mathbf{H}t}\, |\vec{x}_0\rangle &= \frac{m^3}{2\pi^2 r}\int_0^\infty dv\, v\sin(vr)e^{-i\tau\sqrt{1+v^2}} & (12)\\
&= -\frac{m^3}{2\pi^2 r}\frac{d}{dr}\int_0^\infty dv\cos(vr)e^{-i\tau\sqrt{1+v^2}} & (13)
\end{aligned}$$

Gradshtein and Ryzhik 3.914 tells us that the last integral is on the border of the valid region and should be $\dfrac{i\tau}{\sqrt{r^2-\tau^2}}K_1(\sqrt{r^2-\tau^2})$.

It is pleasing to find that, except for the non-covariant normalization factor, the dependence is on the invariant $m^2((\vec{x}-\vec{x}_0)^2-t^2)$ but it is not so pleasant to see that the function does not vanish for positive, spacelike, arguments. Evaluation by steepest descent[4] shows that for $r \gg \tau$ the propagator goes like $e^{-\sqrt{r^2-\tau^2}}$, so it falls exponentially but is not identically zero, as relativity should have it. Thus we have a finite probability that the particle has moved to a position it only could have gotten to by moving faster than the speed of light!

We will see that the resolution of this contains the creation of particle — antiparticle pairs, and the probability of finding a particle at a distance $x > ct$ from where one particle was is due to creation of another particle from the vacuum, followed by the antiparticle member of the pair later annihilating our original particle.




So we see that the quantum mechanics of a single particle gives unphysical results in a relativistic theory, and we need to consider rather quantizing the field.

# 4 Field Quantization

Now consider again the Klein Gordon field described by the "coordinates" $\phi(\vec{x}, t)$ and the canonical momenta $\pi(\vec{x}, t)$, with the Hamiltonian density

$$\mathcal{H} = \frac{1}{2}\pi^2 + \frac{1}{2}\left(\nabla\phi\right)^2 + \frac{1}{2}m^2\phi^2.$$

How do we quantize this? If we go back to the lattice, we would say

$$[P_i, q_j] = -i\delta_{ij}.$$

Our momenta $\pi$, however, are momentum densities, so the normalization we should expect is $\int d^3x[\pi(\vec{x}), \phi(\vec{x}\,')] = -i$ provided $\vec{x}\,'$ is in the integration region, and 0 otherwise. This is what a Dirac delta was invented for, so

$$[\pi(\vec{x}), \phi(\vec{x}\,')] = -i\delta^3(\vec{x} - \vec{x}\,').$$

Of course the coordinates commute with each other, as do the momenta:

$$[\phi(\vec{x}), \phi(\vec{x}\,')] = 0, \qquad [\pi(\vec{x}), \pi(\vec{x}\,')] = 0.$$

[4]See, for example, Arfken, "Mathematical Methods for Physicists", 2nd Ed., pp 373-376.

The classical mechanics of the wave equation is very simple, with solutions corresponding to all 3-momenta $\vec{k}$, with

$$\phi(\vec{x},t) \propto e^{i\vec{p}\cdot\vec{x}-i\omega t}, \quad \text{with } \omega = \pm\sqrt{p^2+m^2},$$

so more generally the field can be expanded in terms of coefficients for each momentum,

$$\phi(\vec{x},t) = \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}\cdot\vec{x}-i\omega t}\,\tilde{\phi}(\vec{p}).$$

Note that as $\phi(\vec{x})$ is real (or, in quantum mechanics, hermitean), $\tilde{\phi}$ is **not real** but rather satisfies the condition $\phi^\dagger(\vec{p}) = \phi(-\vec{p})$. Thus we would do better to write

$$\phi(\vec{x},t) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{p}}}} \left( a_{\vec{p}}\, e^{i\vec{p}\cdot\vec{x}-i\omega_{\vec{p}}t} + a^\dagger_{\vec{p}}\, e^{-i\vec{p}\cdot\vec{x}+i\omega_{\vec{p}}t} \right),$$

which is automatically Hermitean. The factor $1/\sqrt{2\omega_{\vec{p}}}$ is introduced to set the scale of $a$ and $a^\dagger$ conveniently, as we will explain in a minute. From $\pi = \dot{\phi}$ we gather that

$$\pi(\vec{x},t) = -i\int \frac{d^3p}{(2\pi)^3} \sqrt{\frac{\omega_{\vec{p}}}{2}} \left( a_{\vec{p}}\, e^{i\vec{p}\cdot\vec{x}-i\omega_{\vec{p}}t} - a^\dagger_{\vec{p}}\, e^{-i\vec{p}\cdot\vec{x}+i\omega_{\vec{p}}t} \right),$$

At time $t = 0$ it may be more convenient, by changing the dummy variable $\vec{p}$ into $-\vec{p}$ in the $a^\dagger$ terms, to write these as

$$\phi(\vec{x},0) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{p}}}} \left( a_{\vec{p}} + a^\dagger_{-\vec{p}} \right) e^{i\vec{p}\cdot\vec{x}},$$

$$\pi(\vec{x},0) = -i\int \frac{d^3p}{(2\pi)^3} \sqrt{\frac{\omega_{\vec{p}}}{2}} \left( a_{\vec{p}} - a^\dagger_{-\vec{p}} \right) e^{i\vec{p}\cdot\vec{x}},$$

Then

$$\begin{aligned}\begin{pmatrix} [\phi(\vec{x}),\phi(\vec{x}\,')] \\ [\phi(\vec{x}),\pi(\vec{x}\,')] \\ [\pi(\vec{x}),\pi(\vec{x}\,')] \end{pmatrix} &= \begin{pmatrix} 0 \\ i\delta^3(\vec{x}-\vec{x}\,') \\ 0 \end{pmatrix} = \int \frac{d^3p}{(2\pi)^3} \int \frac{d^3p'}{(2\pi)^3} \frac{1}{2\sqrt{\omega_{\vec{p}}\omega_{\vec{p}'}}} e^{i\vec{p}\cdot\vec{x}+i\vec{p}\,'\cdot\vec{x}\,'} \\ &\qquad \times \begin{pmatrix} [a_{\vec{p}} + a^\dagger_{-\vec{p}}, a_{\vec{p}'} + a^\dagger_{-\vec{p}'}] \\ -i\omega_{\vec{p}'}[a_{\vec{p}} + a^\dagger_{-\vec{p}}, a_{\vec{p}'} - a^\dagger_{-\vec{p}'}] \\ -\omega_{\vec{p}}\omega_{\vec{p}'}[a_{\vec{p}} - a^\dagger_{-\vec{p}}, a_{\vec{p}'} - a^\dagger_{-\vec{p}'}] \end{pmatrix}\end{aligned}$$

From the double Fourier transform of this, we see that

$$\begin{aligned}
[a_{\vec p} + a^\dagger_{-\vec p}, a_{\vec p'} + a^\dagger_{-\vec p'}] &= 0 \\
[a_{\vec p} + a^\dagger_{-\vec p}, a_{\vec p'} - a^\dagger_{-\vec p'}] &= -2(2\pi)^3\delta^3(\vec p + \vec p\,') \\
[a_{\vec p} - a^\dagger_{-\vec p}, a_{\vec p'} - a^\dagger_{-\vec p'}] &= 0
\end{aligned}$$

or

$$[a_{\vec p}, a_{\vec p'}] = 0, \qquad [a_{\vec p}, a^\dagger_{\vec p'}] = (2\pi)^3\delta^3(\vec p - \vec p\,'), \qquad [a^\dagger_{\vec p}, a^\dagger_{\vec p'}] = 0.$$

Thus we have commuting operators for different $\vec p$, and for each $\vec p$ we have a set of ordinary harmonic oscillator raising and lowering operators, or at least we would have if we were dealing with momenta discretized by working in a finite box.

The Hamiltonian (at $t = 0$) can be written

$$\begin{aligned}
H &= \int d^3x \left\{ \frac{1}{2}\pi^2(\vec x) + \frac{1}{2}\left(\nabla\phi\right)^2(\vec x) + \frac{1}{2}m^2\phi^2(\vec x) \right\} \\
&= \frac{1}{2}\int d^3x \int \frac{d^3p}{(2\pi)^3}\int \frac{d^3p'}{(2\pi)^3}\Biggl\{ -\frac{\sqrt{\omega_{\vec p}\,\omega_{\vec p'}}}{2}\left(a_{\vec p} - a^\dagger_{-\vec p}\right)\left(a_{\vec p'} - a^\dagger_{-\vec p'}\right) \\
&\qquad + \frac{m^2 - \vec p\cdot\vec p\,'}{2\sqrt{\omega_{\vec p}\,\omega_{\vec p'}}}\left(a_{\vec p} + a^\dagger_{-\vec p}\right)\left(a_{\vec p'} + a^\dagger_{-\vec p'}\right)\Biggr\} e^{i(\vec p\,' + \vec p)\cdot\vec x} \\
&= \frac{1}{2}\int \frac{d^3p}{(2\pi)^3}\Biggl\{ -\frac{\omega_{\vec p}}{2}\left(a_{\vec p} - a^\dagger_{-\vec p}\right)\left(a_{-\vec p} - a^\dagger_{\vec p}\right) \\
&\qquad + \frac{m^2 + \vec p^{\,2}}{2\omega_{\vec p}}\left(a_{\vec p} + a^\dagger_{-\vec p}\right)\left(a_{-\vec p} + a^\dagger_{\vec p}\right)\Biggr\} \\
&= \frac{1}{2}\int \frac{d^3p}{(2\pi)^3}\omega_{\vec p}\Biggl\{ a_{\vec p}\,a^\dagger_{\vec p} + a^\dagger_{\vec p}\,a_{\vec p}\Biggr\} \\
&= \int \frac{d^3p}{(2\pi)^3}\omega_{\vec p}\left\{ a^\dagger_{\vec p}\,a_{\vec p} + \frac{1}{2}[a_{\vec p}, a^\dagger_{\vec p}]\right\} \\
&= \int \frac{d^3p}{(2\pi)^3}\omega_{\vec p}\left\{ a^\dagger_{\vec p}\,a_{\vec p} + \frac{1}{2}\delta^3(0)\right\}
\end{aligned}$$

Notice that the Hamiltonian separates, each spatial momentum component $\vec p$ decoupling from the others and entering as a simple harmonic oscillator. However it is a bit disconcerting to have the zero point energy, $\frac{1}{2}\omega_{\vec p}$, for all of the infinite number of normal modes. As long as we avoid general relativity, and the coupling of this energy to gravitation, we can ignore this constant, though infinite, contribution to the energy of every state in the system — only energy differences have any effect. So we will drop this constant and write

$$H = \int \frac{d^3p}{(2\pi)^3} \omega_{\vec{p}}\, a^\dagger_{\vec{p}}\, a_{\vec{p}}.$$

We can find the possible states of the field theory by examining the commution relations of $H$ with $a_{\vec{p}}$ and $a^\dagger_{\vec{p}}$. As

$$[a^\dagger_{\vec{p}\,'}\, a_{\vec{p}\,'}, a_{\vec{p}}] = [a^\dagger_{\vec{p}\,'}, a_{\vec{p}}]\, a_{\vec{p}\,'} = -(2\pi)^3 \delta^3(\vec{p}\,' - \vec{p}\,)\, a_{\vec{p}},$$

and

$$[a^\dagger_{\vec{p}\,'}\, a_{\vec{p}\,'}, a^\dagger_{\vec{p}}] = a^\dagger_{\vec{p}\,'} [a_{\vec{p}\,'}, a^\dagger_{\vec{p}}] = (2\pi)^3 \delta^3(\vec{p}\,' - \vec{p}\,)\, a^\dagger_{\vec{p}},$$

we have

$$[H, a_{\vec{p}}\,] = \int \frac{d^3p'}{(2\pi)^3} \omega_{\vec{p}\,'}\, [a^\dagger_{\vec{p}\,'}\, a_{\vec{p}\,'}, a_{\vec{p}}\,] = -\omega_{\vec{p}}\, a_{\vec{p}},$$

$$[H, a^\dagger_{\vec{p}}\,] = \int \frac{d^3p'}{(2\pi)^3}\, \omega_{\vec{p}\,'}\, [a^\dagger_{\vec{p}\,'}\, a_{\vec{p}\,'}, a^\dagger\,] = \omega_{\vec{p}}\, a^\dagger_{\vec{p}}.$$

If $|\psi\rangle$ is a state with energy $E$, ($H\,|\psi\rangle = E\,|\psi\rangle$), then $a_{\vec{p}}\,|\psi\rangle$ is a state with energy $E - \omega_{\vec{p}}$, unless $a_{\vec{p}}\,|\psi\rangle = 0$. Thus the lowest energy state of the system must be the $|\psi\rangle$ for which $a_{\vec{p}}\,|\psi\rangle = 0$ for all $\vec{p}$. This is called the vacuum state, and has energy zero (after having dropped the constant term from $H$). Thus we will describe this state as $|0\rangle$, with $H\,|0\rangle = 0$, $a_{\vec{p}}\,|0\rangle = 0$ for all $\vec{p}$.

We can create other states from the vacuum state by applying raising operators $a^\dagger_{\vec{p}}$, repeatedly, to the vacuum state. From their commutator with $H$ we see that the energy is

$$\begin{aligned} H a^\dagger_{\vec{p}_1} a^\dagger_{\vec{p}_2}\,|0\rangle &= [H, a^\dagger_{\vec{p}_1}] a^\dagger_{\vec{p}_2}\,|0\rangle + a^\dagger_{\vec{p}_1}[H, a^\dagger_{\vec{p}_2}]\,|0\rangle + a^\dagger_{\vec{p}_1} a^\dagger_{\vec{p}_2} H\,|0\rangle \\ &= \omega_{\vec{p}_1} a^\dagger_{\vec{p}_1} a^\dagger_{\vec{p}_2}\,|0\rangle + a^\dagger_{\vec{p}_1} \omega_{\vec{p}_2} a^\dagger_{\vec{p}_2}]\,|0\rangle + 0 \\ &= (\omega_{\vec{p}_1} + \omega_{\vec{p}_2}) a^\dagger_{\vec{p}_1} a^\dagger_{\vec{p}_2}\,|0\rangle \end{aligned}$$

so the state $a^\dagger_{\vec{p}_1} a^\dagger_{\vec{p}_2}\,|0\rangle$ has energy (is an eigenstate of $H$ with eigenvalue) $\omega_{\vec{p}_1} + \omega_{\vec{p}_2}$. This is what a state of two noninteracting particles, each of mass $m$, with momenta $\vec{p}_1$ and $\vec{p}_2$ respectively, should have.

This is surely a hint that

$$\left(\prod_{i=1}^{N} a^\dagger_{\vec{p}_i}\right)|0\rangle$$

is a state of $N$ particles of momenta $\vec{p}_1, \ldots, \vec{p}_N$, and we can get further evidence by asking for the total momentum $\vec{P}$ of this state. The momentum operator may be derived from translation invariance, as any continuous symmetry of a theory corresponds to a conserved quantity, which for translations is the total momentum. We will discuss this relationship, called Noether's theorem, after we make a few more observations about the quantum mechanics of a free scalar field. For the moment, let's just accept that $\vec{P}(t) = -\int d^3x\, \dot{\phi}(\vec{x},t)\,\vec{\nabla}\phi(\vec{x},t)$. Expanding that in terms of $a$ and $a^\dagger$,

$$\begin{aligned}
\vec{P}(0) &= -\int d^3x(-i)\int \frac{d^3p'}{(2\pi)^3}\sqrt{\frac{\omega_{\vec{p}'}}{2}}\left(a_{\vec{p}'} - a^\dagger_{-\vec{p}'}\right)e^{i\vec{p}'\cdot\vec{x}} \\
&\qquad \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{p}}}}\left(a_{\vec{p}} + a^\dagger_{-\vec{p}}\right)(i\vec{p})\,e^{i\vec{p}\cdot\vec{x}} \\
&= -\int \frac{d^3p}{(2\pi)^3}\frac{1}{2}\left(a_{-\vec{p}} - a^\dagger_{\vec{p}}\right)\left(a_{\vec{p}} + a^\dagger_{-\vec{p}}\right)\vec{p} \\
&= \int \frac{d^3p}{(2\pi)^3}\,\vec{p}\,a^\dagger_{\vec{p}}a_{\vec{p}},
\end{aligned}$$

where the integral over $x$ has given us a $(2\pi)^3\delta^3(\vec{p}+\vec{p}\,')$ which we have used to do the $d^3p'$ integral, and in the last line we have dropped from the integrand terms which are antisymmetric under $\vec{p} \leftrightarrow -\vec{p}$ (including $\int d^3p\,\vec{p}\,\delta^3(0)$.)

Notice that as $a_{\vec{p}}|0\rangle = 0$ for all $\vec{p}$, $\vec{P}|0\rangle = 0$. Also, because

$$[\vec{P}, a^\dagger_{\vec{p}}] = \int \frac{d^3p'}{(2\pi)^3}\,\vec{p}\,'a^\dagger_{\vec{p}'}(2\pi)^3\delta^3(\vec{p}-\vec{p}\,') = \vec{p}\,a^\dagger_{\vec{p}},$$

when $\vec{P}$ acts on the state $a^\dagger_{\vec{p}_1}a^\dagger_{\vec{p}_2}|0\rangle$ we get

$$\begin{aligned}
\vec{P}\,a^\dagger_{\vec{p}_1}a^\dagger_{\vec{p}_2}|0\rangle &= [\vec{P}, a^\dagger_{\vec{p}_1}]a^\dagger_{\vec{p}_2}|0\rangle + a^\dagger_{\vec{p}_1}[\vec{P}, a^\dagger_{\vec{p}_2}]|0\rangle + a^\dagger_{\vec{p}_1}a^\dagger_{\vec{p}_2}\vec{P}|0\rangle \\
&= \vec{p}_1 a^\dagger_{\vec{p}_1}a^\dagger_{\vec{p}_2}|0\rangle + a^\dagger_{\vec{p}_1}\vec{p}_2 a^\dagger_{\vec{p}_2}]|0\rangle + 0 \\
&= (\vec{p}_1 + \vec{p}_2)a^\dagger_{\vec{p}_1}a^\dagger_{\vec{p}_2}|0\rangle
\end{aligned}$$

so $a^\dagger_{\vec p_1} a^\dagger_{\vec p_2} |0\rangle$ is an eigenstate of $\vec P$ with total momentum $\vec p_1 + \vec p_2$, and of $H$ with total energy $\omega_{\vec p_1} + \omega_{\vec p_2}$, just as one would expect for a state of two noninteracting particles with momenta $\vec p_1$ and $\vec p_2$ respectively. Thus it is clear that we can construct multiparticle states by applying $a^\dagger$'s to the vacuum state.

The normalization, however, is not what Peskin and Schroeder use, which is instead

$$|\vec p\rangle = \sqrt{2\omega_{\vec p}}\, a^\dagger_{\vec p} |0\rangle\,, \qquad |\vec p_1, \vec p_2\rangle = \sqrt{2\omega_{\vec p_1}}\sqrt{2\omega_{\vec p_2}}\, a^\dagger_{\vec p_1} a^\dagger_{\vec p_2} |0\rangle\,, \ldots$$

for reasons they explain, which I will not repeat. (I will also switch from $\omega_{\vec p}$ to $E_{\vec p}$ to be consistent with them, though I don't see why.) For 1-particle states, this means

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

and

$$1\!\!1_{\text{1-part}} = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\vec p}} |\vec p\rangle \langle \vec p|$$

is the projection operator onto single particle states. This normalization is Lorentz invariant, for the obviously (orthochronous[5] ) Lorentz invariant

$$\begin{aligned}
\int \frac{d^4p}{(2\pi)^3} \delta(p^2 - m^2)\Theta(p^0) &= \int \frac{d^3p}{(2\pi)^3} \int_0^\infty dp^0 \delta\Big((p^0)^2 - E^2_{\vec p}\Big) \\
&= \int \frac{d^3p}{(2\pi)^3} \int_0^\infty \frac{d(p^0)^2}{2p^0} \delta\Big((p^0)^2 - E^2_{\vec p}\Big) \\
&= \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\vec p}}.
\end{aligned}$$

Now that we understand $a_{\vec p}$ and $a^\dagger_{\vec p}$ as operators which annihilate or create one particle of momentum $\vec p$, we see that **the field operator $\phi(\vec x)$ can either create or annihilate a particle at $\vec x$.**

So far we have only considered these operators at a given time, as we generally do in Hamiltonian mechanics, or the Schrödinger picture, where the operators are fixed but the states transform with $H \equiv i\partial/\partial t$. In the Heisenberg picture, the operators are time dependent, with

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

[5]Proper orthochronous Lorentz transformations are those which preserve the future direction of time (orthochronous) and which do not reverse right and left hands (proper). That is, $\Lambda^0{}_0 > 0$ and $\det \Lambda^\mu{}_\nu > 0$.

and similarly for all other operators $\mathcal{O}(t)$, with

$$i\frac{\partial}{\partial t}\mathcal{O} = [\mathcal{O}, H].$$

This gives the equations of motion for the field operators

$$\begin{aligned}\frac{\partial}{\partial t}\phi(\vec{x},t) &= \pi(\vec{x},t),\\ \frac{\partial}{\partial t}\pi(\vec{x},t) &= -\left(-\nabla^2 + m^2\right)\phi(\vec{x},t),\end{aligned}$$

or $\left(\partial_\mu\partial^\mu + m^2\right)\phi(x) = 0$.

There are two worthy things to note[6]:

$$\phi(\vec{x},t) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\left(a_{\vec{p}}\, e^{-ip_\mu x^\mu} + a^\dagger_{\vec{p}}\, e^{ip_\mu x^\mu}\right)\Big|_{p^0=E_{\vec{p}}},$$

and

$$\phi(x^\mu) = e^{iHt-i\vec{P}\cdot\vec{x}}\phi(\vec{0},0)e^{-iHt+i\vec{P}\cdot\vec{x}} = e^{iP_\nu x^\nu}\phi(\vec{0},0)e^{-iP_\rho x^\rho}.$$

Let us return to the question of causality which led us to question the idea of using quantum mechanics for a fixed set of particles and argue that field theory was necessary. To do things a little differently than the book, let's consider the two component free Klein-Gordon field, which, as problem 1 of homework 1 shows, is equivalent to a complex field $\phi$ satisfying $(\partial_\mu\partial^\mu + m^2)\phi = 0$. As there is no reality condition on $\phi$, the coefficients of the negative frequency modes are independent of the positive frequency ones, and we write

$$\phi(\vec{x},t) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\left(b_{\vec{p}}\, e^{-i(Et-\vec{p}\cdot\vec{x})} + a^\dagger_{\vec{p}}\, e^{i(Et-\vec{p}\cdot\vec{x})}\right),$$

and $\phi^\dagger$ is, of course

$$\phi^\dagger(\vec{x},t) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}}\left(a_{\vec{p}}\, e^{-i(Et-\vec{p}\cdot\vec{x})} + b^\dagger_{\vec{p}}\, e^{i(Et-\vec{p}\cdot\vec{x})}\right).$$

---

[6] If $a_i$ and $a_i^\dagger$ satisfy $[a_i, a_j^\dagger] = \delta_{ij}$, $M_{ij}$ is any c-number matrix, and $\mathcal{O}$ any function of $a$ and $a^\dagger$,

$$e^{a^\dagger M a}\mathcal{O}(a, a^\dagger)e^{-a^\dagger M a} = \mathcal{O}(e^{-M}a, a^\dagger e^{M}).$$

As $\phi^\dagger(\vec{x},t)\left|0\right\rangle$ is, up to normalization, a one-"b" particle state localized at $\vec{x}$ at time $t$, we may ask what the probability of finding a one-"b"particle state at $\vec{x}\,'$ at time $t'$ by evaluating

$$\begin{aligned} D(x'^\mu, x^\mu) &:= \left\langle 0\right| \phi(\vec{x}\,',t')\phi^\dagger(\vec{x},t)\left|0\right\rangle \qquad (14)\\ &= \int \frac{d^3p'}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}\,'}}} e^{-i(E_{\vec{p}\,'}t' - \vec{p}\,'\cdot\vec{x}\,')} \\ &\quad\times \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\vec{p}}}} e^{i(E_{\vec{p}}t - \vec{p}\cdot\vec{x})} \quad (2\pi)^3\delta^3(\vec{p}\,'-\vec{p}) \\ &= \int \frac{d^3p}{(2\pi)^3}\,\frac{1}{2E_{\vec{p}}}\, e^{-iE_{\vec{p}}\,(t'-t)} e^{i\vec{p}\cdot(\vec{x}' - \vec{x})}. \qquad (15) \end{aligned}$$

As we saw earlier, the integration measure is invariant under **proper orthochronous** Lorentz transformations, and clearly it is invariant under translations. So we may write $D(x'^\mu, x^\mu) = D(x'^\mu - x^\mu)$ and recognize that it is actually a function only of $s^2 = (t'-t)^2 - (\vec{x}\,'-\vec{x})^2$ and the sign of $t'-t$ in the case that $s^2 > 0$. If $s^2 > 0$, with $t' > t$, we may choose to work in a frame with $\vec{x}\,' = \vec{x}$, $t'-t = s > 0$,

$$\begin{aligned} D(x'^\mu - x^\mu) &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\vec{p}}}\, e^{-iE_{\vec{p}}\,s} \\ &= \frac{4\pi}{(2\pi)^3}\int_0^\infty \frac{p^2\,dp}{2E_p}\, e^{-iE_{\vec{p}}\,s} \\ &= \frac{1}{4\pi^2}\int_m^\infty dE\,\sqrt{E^2-m^2}\, e^{-iEs}. \end{aligned}$$

For large time, that is for large values of $s$, the integral is dominated[7] by the lowest value of $E$, and behaves like $e^{-ims}$, as we might expect because the lowest energy component of the state has energy $m$.

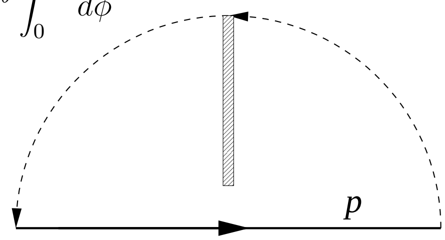
On the other hand, if $x'^\mu - x^\mu$ is spacelike, we can choose to work in a frame with $t' = t$, $\vec{x}\,' - \vec{x} = r\hat{e}_z$, in the $z$ direction[8], and

$$D(x'^\mu - x^\mu) \;=\; \int \frac{d^3p}{(2\pi)^3}\,\frac{1}{2E_{\vec{p}}}\, e^{-ip_z r}$$

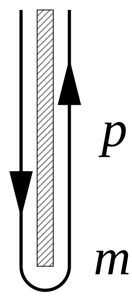
[7]Actually, for real $s$ the integral doesn't actually converge. See footnote 9 for a discussion.

[8]Notation: Although it is $p^\mu$ rather than $p_\mu$ whose spatial components are the usual vector $\vec{p}$, when I use $x$, $y$ and $z$ as in $p_z$, I will write them as subscripts as is customary in Euclidean space. Thus $p_z = p^3$. I hope this dubious notation doesn't cause any difficulty.

$$
\begin{aligned}
&= \frac{1}{(2\pi)^3}\int_0^\infty dp \frac{p^2}{2E_p}\int_0^\pi \sin\theta d\theta e^{-ipr\cos\theta}\int_0^{2\pi} d\phi \\
&= \frac{1}{(2\pi)^2}\int_0^\infty dp \frac{p^2}{2E_p}\frac{e^{ipr}-e^{-ipr}}{ipr} \\
&= \frac{-i}{8\pi^2 r}\int_{-\infty}^\infty dp \frac{p}{\sqrt{p^2+m^2}} e^{ipr}
\end{aligned}
$$




To evaluate this integral[9] we note that for $r > 0$ we may add a semicircle in the upper half complex $p$ plane at $|p| = \infty$ to close the contour. Then we may deform the contour to surround the cut along $p = iy$ for $y \geq m$. Thus

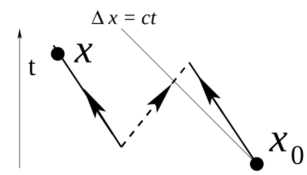



$$
\begin{aligned}
D(x'^\mu - x^\mu) &= \frac{-i}{8\pi^2 r}\int_m^\infty -i\, dy \frac{iy}{-i\sqrt{y^2-m^2}} e^{-yr} \\
&\quad + \frac{-i}{8\pi^2 r}\int_m^\infty i\, dy \frac{iy}{i\sqrt{y^2-m^2}} e^{-yr} \\
&= \frac{-i}{4\pi^2 r}\int_m^\infty i\, dy \frac{iy}{i\sqrt{y^2-m^2}} e^{-yr}.
\end{aligned}
$$

While this does decrease exponentially (roughly as $e^{-mr}$ for large $r$), it is certainly not zero (the integrand is positive everywhere), so we do not have a zero probability of finding an "b" particle at a space-like separated point.

[9] Again we see that the integral doesn't quite converge. The oscillations which cause this are thrown away in this argument at the ends of the arc at infinity. The justification for treating these propagators as if they were well defined is to expect them to be convolved with some wave packet which will smear out the oscillations. Another way to say this is that the oscillations are an indication of delta functions which are irrelevant for nonzero distance separations.

We see that $\langle 0|\, \phi(\vec{x}\,', t')\phi^\dagger(\vec{x}, t)\, |0\rangle$ cannot represent the amplitude for creating a particle from the vacuum at $(\vec{x}, t)$ and have it propagate to $(\vec{x}\,', t')$, but rather something more complicated, including virtual pair creation and annihilations between $t$ and $t'$. In fact, we don't even get zero if $t'$ is earlier than $t$. As we will see later, causality in field theory is related to the *commutator* of operators. So we will now ask if the commutator of two fields vanishes for spacelike separations.




For free fields it is clear that the commutator of two fields at any two times is a c-number, and so we can evalate it by taking its vacuum-expectation value (VEV):

$$[\phi^\dagger(x), \phi(y)] = \langle 0|\, \phi^\dagger(x)\phi(y)\, |0\rangle - \langle 0|\, \phi(y)\phi^\dagger(x)\, |0\rangle = D(x-y) - D(y-x),$$

were I have used the fact that the calculation of $\langle 0|\, \phi(y)\phi^\dagger(x)\, |0\rangle$ is identical to that for $\langle 0|\, \phi^\dagger(y)\phi(x)\, |0\rangle$, only replacing $a$ and $a^\dagger$ with $b$ and $b^\dagger$. For spacelike separations, however, $x-y$ and $y-x$ can be rotated into each other, and as $D$ is invariant under orthochronous Lorentz transformations, which include rotations, the difference must vanish. Thus $[\phi^\dagger(x), \phi(y)] = 0$ for spacelike $x-y$. Of course for the complex field $[\phi(x), \phi(y)] = 0$ for all $x$ and $y$, as the $a$'s commute with both $b$ and $b^\dagger$. We do, however, need to be careful about $[\phi^\dagger(x), \phi(y)] = 0$ near $x^\mu = y^\mu$. For from the lagrangian derived in your homework (with $\phi := (\phi_1 + i\phi_2)/\sqrt{2}$)

$$\mathcal{L} = \partial_\mu\phi^\dagger\partial^\mu\phi - \phi^\dagger\phi,$$

we see that

$$\pi(x) := \frac{\delta\mathcal{L}}{\delta\dot{\phi}(x)} = \dot{\phi}^\dagger(x),$$

and as we know that at equal times

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

we see that

$$\frac{\partial}{\partial t}[\phi^\dagger(\vec{x}, t), \phi(\vec{y}, t')]\bigg|_{t'=t} = -i\delta^3(\vec{y} - \vec{x}).$$

So we see that we have singular behavior for $x^\mu \to y^\mu$.

More generally, for timelike separations we cannot map $x-y$ to $y-x$ with an orthochronous Lorentz transformation, and in fact $D(x-y) = D^*(y-x)$. The commutator is

$$\langle 0|\,[\phi^\dagger(x),\phi(y)]\,|0\rangle = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\vec p}}\left(e^{-ip_\mu(x-y)^\mu} - e^{ip_\mu(x-y)^\mu}\right)\Big|_{p_0=E_p}$$

$$= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\vec p}}\left(e^{-ip_\mu(x-y)^\mu}\Big|_{p_0=E_p} - e^{-ip_\mu(x-y)^\mu}\Big|_{p_0=-E_p}\right) \quad (16)$$

$$= \int \frac{d^4p}{(2\pi)^3} e^{-ip_\mu(x-y)^\mu}\delta(p^2-m^2)\epsilon(p^0) \quad (17)$$

I used the fact that the integration is symmetric under $\vec p$ to change the sign of $\vec x - \vec y$ in writing (16), and the property of the delta function that

$$\delta(f(x)) = \sum_{x_i \ni f(x_i)=0} \frac{\delta(x-x_i)}{|f'(x_i)|}$$

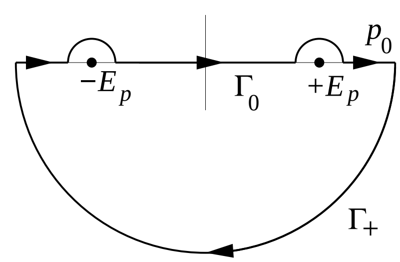
to write (17). $\epsilon(x) = +1$ for $x>0$ and $-1$ for $x<0$.

We can think of the $p_0$ integral

$$\int dp_0 f(p_0)\delta(p_0^2-E^2)\epsilon(p^0)$$

(with $E = \sqrt{\vec p^{\,2}+m^2}$) as the integral around the poles of $\dfrac{1}{2\pi i}\dfrac{f}{p_0^2-E^2}dp_0$.

If $x^0 > y^0$, the integral $\int_{-\infty}^{\infty} dp_0$, distorted as in contour $\Gamma_0$, may be closed in the lower half plane with an infinite semicircle, and the delta function contributions can be considered as residues of poles of $(p^2-m^2)^{-1}$. That is, because




$$\operatorname*{Res}_{p_0=E}\frac{1}{p^2-m^2} = \frac{1}{2E}, \qquad \operatorname*{Res}_{p_0=-E}\frac{1}{p^2-m^2} = -\frac{1}{2E},$$

we can write the commutator as

$$\langle 0|\,[\phi^\dagger(x),\phi(y)]\,|0\rangle = \int \frac{d^3p}{(2\pi)^3}\int_{\Gamma_0}\frac{dp_0}{2\pi i}\frac{-1}{p^2-m^2}\,e^{-ip_\mu(x-y)^\mu} \quad \text{for } x^0>y^0$$

The minus sign is because our contour runs clockwise rather than counter-clockwise.

For $x^0 < y^0$, we can close the contour $\Gamma_0$ with a semicircle $\Gamma_-$ in the upper half plane. This closed contour contains no poles or other singularities, so the integral over $\Gamma_0 \bigcup \Gamma_-$ vanishes. Our function, integrated over $\Gamma_0$, is known as the *retarded* Green's function

$$\begin{aligned} D_R(x-y) &:= \theta(x^0-y^0)\,\langle 0|\,[\phi^\dagger(x),\phi(y)]\,|0\rangle \\ &= \int \frac{d^3p}{(2\pi)^3}\int_{\Gamma_0}\frac{dp_0}{2\pi i}\,\frac{-1}{p^2-m^2}\,e^{-ip_\mu(x-y)^\mu}. \end{aligned}$$

As implied by the "Green's function" part of the name,

$$(\partial_x^2+m^2)D_R(x-y) = -i\delta^4(x-y),$$

where $\partial_x^2 = \partial_\mu\partial^\mu$ involves differentiation with respect to $x$ only, holding $y$ fixed. In the derivation in the book, one needs that the derivative of the heaviside function $\Theta(t)$ is $\delta(t)$, and the derivative of $\delta(t)$ can only be evaluated by differentiating whatever multiplies it,

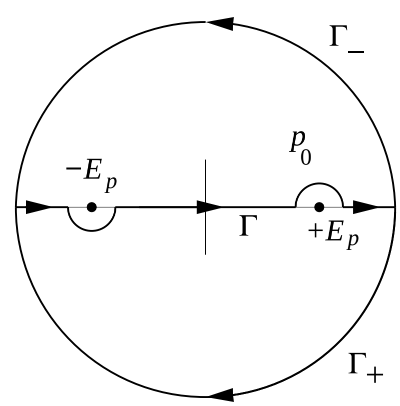
$$\delta'(x)f(x) \sim -\delta(x)f'(x).$$

$D_R(x-y)$ may also be written as the Fourier transform of $\tilde{D}_R(p)$,

$$D_R(x-y) = \int_{\Gamma_0}\frac{d^4p}{(2\pi)^4}\tilde{D}_R(p)\,e^{-ip_\mu(x-y)^\mu}, \quad \text{with } \tilde{D}_R(p) = \frac{i}{p^2-m^2}, \tag{18}$$

with the contour $\Gamma_0$ indicating the contour for $p^0$ for each fixed value of $\vec{p}$.

As with any Green's function, there is an ambiguity due to adding in solutions of the homogeneous equation. In terms of the Fourier transforms, this ambiguity corresponds to choices of how the $p^0$ contour deviates around the poles. An important choice is called the Feynman propagator, defined the same way $D_R$ is in (18), except that the contour is defined to pass below the pole at $p^0 = -E_p$ and above the one at $p^0 = -E_p$, as $\Gamma$ shown.




This choice is sometimes indicated by taking the Fourier transform to be $i/(p^2-m^2+i\epsilon)$, which has the poles at $p_0 = \pm\sqrt{E_{\vec{p}}^2 - i\epsilon} = \pm(E_{\vec{p}} - i\epsilon')$ for some infinitesimal but positive values of $\epsilon$ and $\epsilon'$. This places the poles just above the negative real axis and just below the positive one, so $dp_0$ can run exactly on the real axis and pass to the sides of the poles Feynman requested. Closing the contour with $\Gamma_+$ in the lower half plane for $x^0 > y^0$ or with $\Gamma_-$ in the upper half plane for $x^0 < y^0$, we see that only the pole at $p^0 = E_{\vec{p}}$ contributes for $x^0 > y^0$, and that at $p^0 = -E_{\vec{p}}$ for $x^0 < y^0$. The residues of $(p^2-m^2)^{-1}$ are $+1/2E_p$ and $-1/2E_p$ respectively, but the clockwise direction of the contour $\Gamma \bigcup \Gamma_+$ gives an extra $-$ sign in the $x^0 > y^0$ case. Then we have

$$\begin{aligned} D_F(x-y) &= \theta(x^0-y^0)D(x-y) + \theta(y^0-x^0)D(y-x) \\ &= \theta(x^0-y^0)\left\langle 0\right| \phi^\dagger(x)\phi(y)\left|0\right\rangle + \theta(y^0-x^0)\left\langle 0\right| \phi(y)\phi^\dagger(x)\left|0\right\rangle \\ &=: \left\langle 0\right| T\phi^\dagger(x)\phi(y)\left|0\right\rangle . \end{aligned} \tag{19}$$

In the last expression, $T$ is a kind of meta "time-ordering" operator, which tells us to rewrite the operators following it in chronological sequence (right to left increasing).

Thus far we have discussed only a free field — that is, as we have seen, the solution to the equations of motion are all of the form that we have some bunch of particles of various momenta, but each one has constant momentum, and the number of particles is unchanged in time. This is elegant but not very interesting — we need to include interactions to have interesting physics. Unfortunately, interactions make the theory much more difficult to solve. The simplest interaction our Klein-Gordon particles can have is with an external classical source. In electromagnetism, this classical source is a 4-current $j^\mu$ as you considered for homework, which enters as an additional term $-A_\mu j^\mu$ in the lagrangian. There are complications here, but for a real scalar field the source is simpler, so that

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}m^2\phi^2 + j(x)\phi(x) \quad\Longrightarrow\quad (\partial^2+m^2)\phi = j(x).$$

Thus the field $\phi(x)$ can be found using the retarded Green function

$$\phi(x) = \phi_0 + i\int d^4y\, D_R(x-y)j(y).$$

The last section of Chapter 2 shows that the source creates particles of momentum $\vec{p}$ with a number expectation value given by the square of the fourier transform

$$\tilde{j}(\sqrt{\vec{p}^{\,2}+m^2},\vec{p}) = \int d^4y\, e^{ip\cdot y} j(y)\Big|_{p^0=\sqrt{\vec{p}^{\,2}+m^2}} .$$

You should read through this section, but I will not lecture on it.