Quantum Theory

Dan Wohns

This text was prepared for the 2023/2024 Quantum Theory course and is solely for the purpose of studying in the course and is not meant for further distribution.

Recommended reading:

- M. D. Schwartz, *Quantum Field Theory and the Standard Model*
- M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory*
- A. Zee, *Quantum Field Theory in a Nutshell*
- D. Tong, Lecture notes on quantum field theory, http://www.damtp.cam.ac.uk/user/tong/qft.html
- R. Casalbuoni, *Introduction to Quantum Field Theory*

Additional sources and references:

- S. Weinberg, *Quantum Theory of Fields I*
- M. Srednicki, *Quantum Field Theory*
- A. Duncan, *The Conceptual Framework of Quantum Field Theory*
- T. Lancaster and S.J. Blundell, *Quantum Field Theory for the Gifted Amateur*
- W. Greiner and J. Reinhart, *Field Quantization*
- F. David, Lecture notes
- T. Ali, Lecture notes

July 31, 2023

Contents

1. The Problem - Collision of Two Particles

What happens after two objects collide?

This question will be the primary focus of the remainder of this course.[1] In some limits, you already know how to predict what occurs after a collision.

If the objects are large and slow, Newtonian mechanics allows us to predict the trajectories of the objects after the collision. For large, fast objects, relativistic corrections to Newtonian mechanics become important so we must use special relativity.

If the objects are small enough, quantum mechanics predicts the probabilities of the allowed possible final states of the two objects. Non-relativistic quantum mechanics fails if the objects are sufficiently fast. One reason for this failure is that non-relativistic quantum mechanics describes the time evolution of some fixed set of objects or particles. The initial and final state contain the same particles so particle number is conserved.

There is a simple argument that suggests when we combine quantum mechanics and special relativity we should not expect particle number to be conserved. The uncertainty principle gives a lower bound to the uncertainty in momentum[2] $\Delta p \geq \frac{1}{L}$ for a particle in a box of size L. Special relativity says that momentum, energy, and mass are all related $p \approx E = m$. If the box is small enough so that $\Delta E \geq 2m$ particle-antiparticle pairs with mass m can be created. Any attempt to make single-particle[3] quantum mechanics relativistically invariant will fail.

To describe collision in which the number or type of particles change, we need to combine quantum mechanics and special relativity. For reasons described in the next section, this combination of theories is known as quantum field theory.

	Large	Small
Slow	Newtonian mechanics	Quantum mechanics
Fast	Special relativity	Quantum field theory

In addressing the question of what happens after objects collide, we will need to think carefully about what quantities can be measured. The questions we can ask about the collision process depends on which limit we are considering. For example, we can ask about the trajectory of each of the particles. In the classical limit where the colliding objects are large, this question is sensible and has a unique answer. When the objects are small enough, many trajectories contribute. A question that often has a unique and well-defined answer in the classical limit does not make sense quantum mechanically (or at least does not have a unique answer).

[1] We will take a few small detours. Many of the concepts and tools we will develop to address scattering problems have other applications.

[2] Throughout the course we will use natural units $\hbar = c = 1$. If you prefer to use unnatural units, you can always restore the $\hbar$'s and c's using dimensional analysis.

[3] Or n-particle quantum mechanics for any fixed n.

Quantum field theory is a theory of nature. A given quantum field theory makes finite, unique predictions for any measurable physical quantity. One way to specify a quantum field theory is by the results of a finite set of measurements. This finite set of measurements might include the masses of the particles in the theory as well as the amplitude for some particular scattering process at a particular energy. Other observables, including amplitudes measured at other energies, can then be computed. In this sense measurements at all energy scales[4] can be predicted once measurements at a reference energy scale are known. This process is known as renormalization. We will briefly discuss in the last part of this course. You will learn more about renormalization in Quantum Field Theory I and II and in Statistical Physics.

In Newtonian physics, many calculations involve non-observable quantities. For example the easiest solution to a problem may involve a choice of origin for a coordinate system. As the coordinate system is not physical, none of the predictions of Newtonian mechanics depend on this choice. Special relativity, quantum mechanics, and quantum field theory computations also involve non-observable quantities. Unlike in classical physics or non-relativistic quantum mechanics, in quantum field theory many non-observable quantities that appear in intermediate steps in computations are infinite. Physical measurements are always finite and the infinities we will encounter will cancel out in any quantum field theory calculation of a quantity that can be measured in a physical experiment.

Before we can answer the question of what happens when small fast objects collide, we first turn to the simpler question of how to describe a relativistic quantum particle in the absence of interactions.

2. Why Quantum Field Theory (QFT)?

Quantum mechanics and special relativity are two of the most successful theories of nature. In this course we will combine these two theories. To do so we are essentially required to use fields.

We can certainly try to combine quantum mechanics and special relativity without using fields. In tutorial we will compute the amplitude for a relativistic particle at the origin ($t = 0$, $\mathbf{x} = 0$) to travel faster than the speed of light to some position (t, $\mathbf{x}$) in single-particle quantum mechanics. We want this amplitude to be zero so that the probability of faster-than-light travel is zero. You will show that it is non-zero.

We could build quantum field theory starting from first principles. Quantum field theory seems to be the only way to combine quantum mechanics with special relativity with the additional assumption of clustering. Clustering means that local experiments are unaffected by the distant environment[5]. You will see this approach in Quantum Field Theory I. In this

[4] At least for energies where the theory is valid. Typically quantum field theories are valid up to some maximum energy scale.

[5] More precisely, the cluster decomposition property requires that the vacuum expectation value of localized

approach after constructing a quantum field theory classical field theory can be recovered by taking a suitable limit.

In this course we will take the opposite approach. We will start with a classical field theory and then quantize it by promoting the fields to field operators. We will see that the resulting quantum field theory is consistent with quantum mechanics, special relativity, and clustering.

At first glance it may not be surprising that quantizing a Lorentz-invariant quantum field theory results in a quantum field theory that is consistent with special relativity and clustering. Before quantizing our theory will obey special relativity. One of the reasons to introduce classical fields in the first place is to avoid action at a distance. For example the electromagnetic field explains locally how distant charges interact. These properties are preserved after quantization even though not all steps in our quantization procedure will be manifestly Lorentz invariant. In the end we will recover a Lorentz-invariant quantum field theory.

Quantum field theory explains several features that are mysterious in non-relativistic quantum mechanics:

- First quantum field theory explains why identical particles are identical. For example all electrons have exactly the same mass and charge. In non-relativistic quantum mechanics there is no reason why the masses and charges need to be the same. In quantum field theory identical particles are excitations of a single field defined over all of spacetime.

- Quantum field theory also explains the existence of antimatter. In quantum field theory interactions between particles occurs through local interactions with other particles. Proton-neutron scattering can occur through the exchange of a virtual charged pion. If the interaction vertices are spacelike separated, then in different frames the charge of the pion should change. This observation suggests that the existence of a positively charged implies the existence of a negatively charged pion, or anti-pion.

- Additionally the spin-statistics theorem explains why integer spin particles obey Bose-Einstein statistics and half-integer spin particles obey Fermi-Dirac statistics. In non-relativistic quantum mechanics the relationship between spin and statistics must be assumed. In quantum field theory this relationship is a consequence of Lorentz invariance.

There are two[6] possible equivalent approaches we could take to quantize a classical field theory. In the canonical quantization approach we will use in this course, we will

operators factorize in the limit that the separation between the regions of support of the operators becomes large.

[6] Other quantization methods exist. See "Quantization Methods: A Guide for Physicists and Analysts" by S. Ali and M. Engliš for a discussion of some other methods.

promote the fields to operators. In QFT II, we will use functional integral quantization which is a generalization of the path integral quantization of non-relativistic quantum mechanics. These two methods are equivalent and each has their advantages. For example the connection between fields and particles will be more clear in the canonical formulation. The field operators will be linear combinations of creation and annihilation operators for particles. Some exact results will be easier to demonstrate in the canonical formulation as well. The functional integral approach facilitates the study of non-abelian gauge theories (such the theory of the strong interaction) and instantons (which can be used to study non-perturbative phenomenon such as tunneling) for example.

In quantum mechanics position and time are treated very differently. Position is an operator corresponding to an observable while time is simply a parameter. Treating time and space differently is probably not the best idea if we want to construct a relativistic quantum theory. To treat time and space symmetrically we can either promote time to an operator or demote position to a parameter.

If we promote time to an operator, we still need a time parameter for evolution in the Schrödinger equation. We can use proper time τ as the evolution parameter, and let $\hat{X}^\mu(\tau)$ be the (Heisenberg picture) operators corresponding to coordinate time and position. It is complicated to develop quantum field theory along these lines due to the fact that we could use $f(\tau)$ as our time parameter instead where f is a monotonic function. String theory uses this approach while adding another parameter σ so $\hat{X}^\mu(\tau, \sigma)$ is an operator corresponding to a propagating string.

Instead we will demote position to a label. Field operators (in the Heisenberg picture) can be thought of as a collection of operators at each point in spacetime. Neither position nor time should be thought of as eigenvalues of any hermitian operator in our approach.

3. Classical Field Theory

Our starting point for constructing a quantum field theory will be a classical field theory. One example of a field is the vector potential $A^\mu(t, \mathbf{x})$ from electromagnetism. A^μ transforms as a vector under Lorentz transformations. As you will see in QFT I, there is a connection between the Lorentz transformation properties of a classical field and the spin (or intrinsic angular momentum) of the particles associated to that field after quantization. In this course we will restrict ourselves to scalar fields which transform trivially under Lorentz transformations. The particles associated to scalar fields have spin zero.

A real scalar field $\varphi(t, \mathbf{x}) = \varphi(x)$ has an infinite number of degrees of freedom, one per point in space. Throughout these notes we will indicate four-vectors such as spacetime coordinates x using italic text and three-vectors such as spatial coordinates $\mathbf{x}$ using bold text (except in an integration measure where we distinguish spatial integration measures d^3x from spacetime integration measures d^4x by the exponent).

We will use the mostly minus sign convention for the Minkowski metric

$$\eta^{\mu\nu} = \eta_{\mu\nu} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{1}$$

Occasionally it will be useful to discretize spacetime. On the discretized spacetime, each point of space has one degree of freedom. In the language of classical particle mechanics in the discretized theory we have a finite number of generalized coordinates $q_a(t)$. Classical scalar field theory can be constructed as a continuum limit.

Field theory	Notation	Particle mechanics	Notation
field	φ	position	q
position	$\mathbf{x}$	particle label	a

3.1. Lagrangian Formalism

The Lagrangian formalism is useful for quantum field theory because Lorentz invariance is often manifest.

For example, the action of classical mechanics of a finite set of particles with generalized positions q_a is

$$S = \int dt \sum_a L(q_a, \dot{q}_a)\,. \tag{2}$$

If we think of the generalized positions as the value of a scalar field at some points in spacetime, then in the continuum limit the action of a scalar field we will consider is the integral of a functional[7] of the field and derivatives of the field

$$S = \int dt \int d^3x \mathcal{L}(\varphi(x), \partial_\mu \varphi(x))\,. \tag{3}$$

Here $\mathcal{L}$ is the Lagrangian density, although it is commonly and slightly imprecisely referred to as the Lagrangian. We will restrict our attention to local and Lorentz invariant actions of this form. Locality prevents direct couplings of fields at different spacetime points. An action of the form

$$S = \int dt \int d^3x \int d^3x' \mathcal{L}(\varphi(x), \varphi(x'), \partial_\mu \varphi(x), \partial_\mu \varphi(x')) \tag{4}$$

violates locality. Lorentz invariance requires derivatives show up in covariant combinations. A theory with an action

$$S = \int dt \int d^3x \mathcal{L}(\varphi(x), \dot{\varphi}(x)) \tag{5}$$

[7] A functional is a map from a function space to the real or complex numbers. A functional can be thought of as a function with infinitely many arguments (the value of the function at each point).

is not generally Lorentz invariant.

The equation of motion in classical particle mechanics is given by the variation principle

$$\frac{\delta S}{\delta q_a(t)} = 0 \tag{6}$$

or the Euler-Lagrange equation

$$\frac{\partial L}{\partial q_a} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_a} = 0\,. \tag{7}$$

In field theory, similarly we have

$$\frac{\delta S}{\delta \varphi(x)} = 0\,. \tag{8}$$

If the action is given by (3) the Euler-Lagrange equation is

$$\frac{\partial \mathcal{L}}{\partial \varphi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \right) = 0\,. \tag{9}$$

3.2. Hamiltonian Formalism

Perhaps the most familiar formulation of quantum mechanics uses the Hamiltonian formalism for quantization. In this approach, generalized positions q_a and their conjugate momenta p_a are promoted to operators. We will take the same approach to quantize the fields. Lorentz invariance is not manifest in the Hamiltonian formalism due to the special role that time plays.

We introduce momentum conjugate to $\varphi(x)$:

$$\pi(x) = \frac{\partial \mathcal{L}}{\partial \dot{\varphi}(x)} \tag{10}$$

and the Hamiltonian density:

$$\mathcal{H} = \pi(x)\dot{\varphi}(x) - \mathcal{L}(x)\,. \tag{11}$$

The momentum and Hamiltonian densities are not Lorentz invariant as time and space play different roles in (10) and (11). The Hamiltonian is the integral of the Hamiltonian density

$$H = \int d^3x \mathcal{H}\,. \tag{12}$$

All of the equations in this subsection can be understood by recalling the analogy between the field and generalized position $\varphi \leftrightarrow q$ and position and particle label $\mathbf{x} \leftrightarrow a$.

3.3. Classical Klein-Gordon Theory

The simplest theory is Klein-Gordon theory which has the Lagrangian (recall $\hbar = c = 1$)

$$\mathcal{L}_{\rm KG} = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}m^2\varphi^2 \,. \tag{13}$$

Here the scalar field φ is a real scalar field. As we will see, the quantum version of this theory describes a free neutral spin-0 particle.

The reason that this theory is simple is because the Klein-Gordon equation of motion is linear

$$\partial_\mu\partial^\mu\varphi + m^2\varphi = 0 \tag{14}$$

so the classical theory can be solved exactly. We will see later that the quantum Klein-Gordon theory can also be solved exactly.

The momentum is the time derivative of the field

$$\pi = \dot{\varphi} \,. \tag{15}$$

The Hamiltonian contains a kinetic term, a gradient term, and a potential term

$$H = \int d^3x \left(\frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\varphi)^2 + \frac{1}{2}m^2\varphi^2\right) . \tag{16}$$

As expected, the Hamiltonian is not Lorentz invariant.

3.4. Classical Solution of Klein-Gordon Equation

If we Fourier transform the scalar field (with respect to space only - this approach is slightly easier than a 4d Fourier transform)

$$\varphi(t,\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{x}} \tilde{\varphi}(t,\mathbf{k}) \,, \tag{17}$$

the Klein-Gordon equation (14) becomes

$$(\partial_t^2 + \mathbf{k}^2 + m^2)\tilde{\varphi}(t,\mathbf{k}) = 0 \,. \tag{18}$$

The solution to the Klein-Gordon equation is a linear combination of two exponentials

$$\tilde{\varphi}(t,\mathbf{k}) = A(\mathbf{k})e^{-iE_\mathbf{k}t} + B(\mathbf{k})e^{iE_\mathbf{k}t} \tag{19}$$

where

$$E_\mathbf{k} = \sqrt{\mathbf{k}^2 + m^2} \,. \tag{20}$$

The field is equal to its complex conjugate $\varphi = \varphi^*$ since we are considering a real scalar field. The complex conjugate of φ is

$$\begin{aligned}\varphi^*(t,\mathbf{x}) &= \int \frac{d^3k}{(2\pi)^3} e^{-i\mathbf{k}\cdot\mathbf{x}} \tilde{\varphi}^*(t,\mathbf{k}) \\ &= \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{x}} \tilde{\varphi}^*(t,-\mathbf{k})\end{aligned} \tag{21}$$

where in the second line, we integrate over $-\mathbf{k}$ instead of $\mathbf{k}$. The change of variables $\mathbf{k} \to -\mathbf{k}$ produces an even number of overall minus signs. We get a factors of -1 from the measure and also from the exchange of the integration limits. Rewriting the reality condition $\varphi = \varphi^*$ in terms of the Fourier transform we see that

$$\tilde{\varphi}^*(t,-\mathbf{k}) = \tilde{\varphi}(t,\mathbf{k})\,. \tag{22}$$

The scalar field is real as long as the constraint

$$B(\mathbf{k}) = A^*(-\mathbf{k}) \tag{23}$$

on the mode functions in (19) is satisfied. Hence we have

$$\begin{aligned}\varphi(t,\mathbf{x}) &= \int \frac{d^3k}{(2\pi)^3} \left[A(\mathbf{k}) e^{-iE_{\mathbf{k}}t + i\mathbf{k}\cdot\mathbf{x}} + A^*(-\mathbf{k}) e^{iE_{\mathbf{k}}t + i\mathbf{k}\cdot\mathbf{x}}\right] \\ &= \int \frac{d^3k}{(2\pi)^3} \left[A(\mathbf{k}) e^{-ik\cdot x} + A^*(\mathbf{k}) e^{+ik\cdot x}\right]\end{aligned} \tag{24}$$

where we have used $k^\mu = (E_{\mathbf{k}}, \mathbf{k})$ so that

$$k\cdot x = \eta_{\mu\nu} k^\mu x^\nu = E_{\mathbf{k}} t - \mathbf{k}\cdot\mathbf{x}\,. \tag{25}$$

We also used integration over $-\mathbf{k}$ instead of $\mathbf{k}$ trick again to rewrite the second term as a Lorentz invariant product of four-vectors.

Note that the measure $\int \frac{d^3k}{(2\pi)^3}$ is not Lorentz invariant. Writing our solution in terms of a Lorentz invariant measure will save us some effort later (it will ensure the states created by the creation operators are Lorentz covariant). We can make the measure Lorentz invariant using the following identity involving the δ-distribution (see Appendix A for more on distributions)

$$\int \frac{d^4k}{(2\pi)^4} 2\pi\delta(k^2 - m^2)\Theta(k^0) = \int \frac{d^3k}{(2\pi)^3} \frac{1}{2E_{\mathbf{k}}}\,. \tag{26}$$

The left hand side of the above equation is manifestly Lorentz invariant under orthochronous transformations (transformations that preserve the arrow of time). If we change the normalization of our mode functions

$$A(\mathbf{k}) = \frac{a(\mathbf{k})}{2E_{\mathbf{k}}}\,, \tag{27}$$

we can write the solution in a way that has better Lorentz tranformation properties

$$\varphi(t,\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3 2E_{\mathbf{k}}} \left[a(\mathbf{k})e^{-ik\cdot x} + a^*(\mathbf{k})e^{+ik\cdot x}\right]. \tag{28}$$

This form of the solution to the Klein-Gordon equation will be our starting point when we quantize the field in the next chapter. The mode functions $a(\mathbf{k})$ and $a^*(\mathbf{k})$ will become operators that annihilate or create particles with momentum $\mathbf{k}$.

3.5. Noether's Theorem

Conservation laws are powerful tools for solving classical mechanics problems as well as quantum mechanics problems. Identifying conserved quantities in the scattering of relativistic quantum particles will aid us in our quest to understand scattering.

Noether's theorem tells us every continuous symmetry of the action gives rise to a conserved current $j^\mu(x)$, where conserved means $\partial_\mu j^\mu = 0$.

A conserved current implies a conserved charge

$$Q = \int d^3x j^0 . \tag{29}$$

If we have a continuous symmetry we can consider an infinitesimal transformation

$$\delta\varphi_a(x) = Y_a(\varphi) . \tag{30}$$

We are allowing the possibility that our theory contains multiple scalar fields. Here a is an index that labels the scalar fields. This transformation is a symmetry transformation if the Lagrangian changes by a total derivative

$$\delta\mathcal{L} = \partial_\mu F^\mu(\varphi) \tag{31}$$

for some functions $F^\mu(\varphi)$ since the equations of motion are unchanged.

Let us prove Noether's Theorem. Under an arbitrary transformation of the fields (not necessarily a symmetry), the Lagrangian changes by

$$\begin{aligned}\delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\varphi_a}\delta\varphi_a(x) + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_a)}\partial_\mu(\delta\varphi_a(x)) \\ &= \left[\frac{\partial\mathcal{L}}{\partial\varphi_a} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_a)}\right]\delta\varphi_a(x) + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_a)}\delta\varphi_a(x)\right)\end{aligned} \tag{32}$$

The term in the square brackets vanishes when the equation of motion is satisfied. For a symmetry transformation $\delta\varphi_a = X_a(\varphi)$ the Lagrangian changes by a total derivative $\delta\mathcal{L} = \partial_\mu F^\mu(\varphi)$. Combining these results we see

$$\partial_\mu F^\mu(\varphi) = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_a)}Y_a(\varphi)\right). \tag{33}$$

We can identify the conserved current satisfying $\partial_\mu j^\mu = 0$ from (33)

$$j^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_a)} Y_a(\varphi) - F^\mu(\varphi)\,. \tag{34}$$

Symmetries remain a powerful tool for studying field theories after quantization. We will occasionally make use of Noether's theorem in this course to find the conserved charges associated with symmetries of the field theory we are studying.

Noether's theorem does **not** guarantee that a symmetry of the classical theory implies a conserved current in the quantum theory. Quantum field theories can have anomalies (quantum effects that violate classical symmetries). The theories we will study in this course and QFT I and II will mostly be anomaly free. If a theory is anomaly free the charges and currents identified by Noether's theorem will be conserved.

Example: U(1) internal (global) symmetry.

Consider a complex scalar field Φ. We can think of a complex scalar field as two real fields φ_1 and φ_2

$$\Phi(x) = \frac{1}{\sqrt{2}}(\varphi_1(x) + i\varphi_2(x))\,. \tag{35}$$

Often it is easier to work with Φ and Φ^* instead of φ_1 and φ_2 (as in Wirtinger derivatives where partial derivatives with respect to a complex variable z and its complex conjugate z^* behave essentially like ordinary derivatives with respect to a real variable). Suppose we are interested in the following theory

$$\mathcal{L} = \partial_\mu \Phi^* \partial^\mu \Phi - V(|\Phi|^2)\,. \tag{36}$$

This theory is invariant under the $U(1)$ transformation

$$\Phi \to e^{i\theta}\Phi \tag{37}$$

$$\Phi^* \to e^{-i\theta}\Phi^*\,. \tag{38}$$

The infinitesimal version of the above transformation is

$$\delta\Phi = i\theta\Phi \tag{39}$$

$$\delta\Phi^* = -i\theta\Phi^*\,.$$

Because the transformation does not change the Lagrangian $\delta\mathcal{L} = 0$ there is no total derivative term $F^\mu = 0$ and we have

$$j^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi)}\delta\Phi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi^*)}\delta\Phi^*\,. \tag{40}$$

We can read off the conserved current

$$j^\mu = i\theta(\partial^\mu\Phi^*)\Phi - i\theta\Phi^*(\partial^\mu\Phi)\,. \tag{41}$$

This current is conserved for any θ so the current

$$j^\mu = i(\partial^\mu\Phi^*)\Phi - i\Phi^*(\partial^\mu\Phi) \tag{42}$$

is also conserved. The associated conserved charge is electric charge or particle number.

4. Canonical Quantization of a Free Scalar Field

In tutorial we saw that it is inconsistent to consider a relativistic quantum theory of a single particle. In this chapter we construct a free (non-interacting) relativistic quantum theory of spin-0 particles by quantizing the Klein-Gordon theory. Before we can consider a non-trivial scattering problem, we must develop the machinery to describe the trivial scattering of two relativistic quantum particles that do not interact with each other.

4.1. Canonical Quantization

Our approach for constructing a free scalar quantum field theory will closely mirror the canonical quantization procedure for non-relativistic quantum mechanics. Canonical quantization consists of four steps:

1. Choose a classical, Lorentz-invariant Lagrangian.
2. Compute the conjugate momentum for each field and compute the Hamiltonian.
3. Impose commutation relations between the fields and their conjugate momenta.
4. Impose an ordering prescription.

We will illustrate the canonical quantization procedure using the free Klein-Gordon theory as an example.

Step 1: We choose $\mathcal{L} = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}m^2\varphi^2$ as our Lagrangian. The resulting QFT is guaranteed to be Lorentz-invariant because we started with a Lorentz-invariant Lagrangian.

Step 2: We computed the conjugate momenta and Hamiltonian in the previous chapter in equations (15) and (16). We found

$$\begin{aligned} \pi &= \dot{\varphi} \\ \mathcal{H} &= \frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\varphi)^2 + \frac{1}{2}m^2\varphi^2 \,. \end{aligned} \tag{43}$$

Step 3: In quantum mechanics, the commutation relations are:

$$\begin{aligned} [q_a, p_b] &= i\delta_{ab} \\ [q_a, q_b] &= 0 = [p_a, p_b] \,. \end{aligned} \tag{44}$$

Recalling that we are treating position as a label, the natural generalization of the above commutation relations to quantum field theory is (in the Schrödinger picture)

$$\begin{aligned} [\varphi(\mathbf{x}), \pi(\mathbf{y})] &= i\delta(\mathbf{x}-\mathbf{y}) \\ [\varphi(\mathbf{x}), \varphi(\mathbf{y})] &= 0 = [\pi(\mathbf{x}), \pi(\mathbf{y})] \,. \end{aligned} \tag{45}$$

In section 3.4 we found the classical solution for $\varphi(x)$ in equation (28)

$$\varphi(t,\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3 2E_{\mathbf{k}}} \left[a(\mathbf{k})e^{-ik\cdot x} + a^*(\mathbf{k})e^{+ik\cdot x}\right]. \tag{46}$$

In the last section in this chapter we will introduce the Heisenberg picture field operators that are time dependent. For now we can set $t = 0$ (the Heisenberg and Schrödinger picture operators should agree at some time which we can take to be $t = 0$.) Setting $t = 0$ gives (recall we are using the mostly minus convention for the metric)

$$\varphi(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3 2E_{\mathbf{k}}} \left[a(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}} + a^*(\mathbf{k})e^{-i\mathbf{k}\cdot\mathbf{x}}\right]. \tag{47}$$

After we impose the commutation relations the field becomes a field operator. The right hand side of (47) should be an operator as well. The natural way to accomplish this goal is to promote $a(\mathbf{k})$ and $a^*(\mathbf{k})$ to operators

$$\begin{aligned} a(\mathbf{k}) &\to a_{\mathbf{k}} \\ a^*(\mathbf{k}) &\to a^\dagger_{\mathbf{k}}. \end{aligned} \tag{48}$$

The field operator can be written as a linear combination of creation and annihilation operators

$$\varphi(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3 2E_{\mathbf{k}}} \left[a_{\mathbf{k}}e^{i\mathbf{k}\cdot\mathbf{x}} + a^\dagger_{\mathbf{k}}e^{-i\mathbf{k}\cdot\mathbf{x}}\right]. \tag{49}$$

As you will show in a quiz, the commutation relations (45) imply the following commutation relationships for the operators $a_{\mathbf{k}}$ and $a^\dagger_{\mathbf{k}}$

$$\begin{aligned} [a_{\mathbf{k}}, a^\dagger_{\mathbf{p}}] &= (2\pi)^3 2E_{\mathbf{k}}\delta(\mathbf{k}-\mathbf{p}) \\ [a_{\mathbf{k}}, a_{\mathbf{p}}] &= 0 = [a^\dagger_{\mathbf{k}}, a^\dagger_{\mathbf{p}}]. \end{aligned} \tag{50}$$

Also in a quiz you will show that the Hamiltonian is just the Hamiltonian for infinitely many decoupled harmonic oscillators

$$H = \frac{1}{2}\int \frac{d^3k}{(2\pi)^3 2E_{\mathbf{k}}} E_{\mathbf{k}}(a^\dagger_{\mathbf{k}}a_{\mathbf{k}} + a_{\mathbf{k}}a^\dagger_{\mathbf{k}}). \tag{51}$$

Step 4: The procedure we have described so far is ambiguous. To see why the procedure is ambiguous, consider adding $\varphi\dot{\varphi} - \dot{\varphi}\varphi$ to the classical Klein-Gordon Lagrangian in step 1. Classically $\varphi\dot{\varphi} - \dot{\varphi}\varphi = 0$ so our new classical theory is identical to the classical Klein-Gordon theory. If we follow the canonical quantization procedure above we will find that the Hamiltonian operator is changed by the addition of the new term since in the quantum theory $[\varphi(\mathbf{x}), \pi(\mathbf{y})] = i\delta(\mathbf{x}-\mathbf{y})$. To fix this ambiguity problem we will need an ordering procedure. We will give an explicit ordering procedure in the next section.

The existence of this ordering ambiguity is not unique to quantum field theory. It is present in nonrelativistic quantum mechanics as well. Classical physics is an approximation to quantum physics (in the limit $\hbar \to 0$). Multiple quantum theories can have the same classical limit.

4.2. States

Now that we have constructed the Hamiltonian (up to the ordering issue discussed above) and field operators, we examine the states in the free scalar field theory.

4.2.1. Vacuum

The simplest state is the state with no particles, the vacuum state. The annihilation operators $a_{\mathbf{k}}$ remove a particle when they are applied to a ket. The vacuum state $|0\rangle$ has no particles so the annihilation operators annihilate the vacuum

$$a_{\mathbf{k}}|0\rangle = 0 \tag{52}$$

for any $\mathbf{k}$. We choose to normalize the vacuum as

$$\langle 0|0\rangle = 1\,. \tag{53}$$

Now let us calculate the energy E_0 of the vacuum.

$$\begin{aligned} H|0\rangle &= E_0|0\rangle \\ &= \frac{1}{2}\int \frac{d^3k}{(2\pi)^3 2E_{\mathbf{k}}} E_{\mathbf{k}} a_{\mathbf{k}} a_{\mathbf{k}}^\dagger |0\rangle \\ &= \frac{1}{2}\int d^3k E_{\mathbf{k}} \delta(\mathbf{0})|0\rangle \\ &= \infty|0\rangle\,. \end{aligned} \tag{54}$$

Our first calculation in quantum field theory yields a divergent answer! We will encounter infinities frequently in this course. As all physical measurements are finite, when we encounter an infinity in QFT either we have calculated a quantity that is in principle unobservable or we have made a mistake somewhere. Let us examine the origin of this divergence more closely.

There are two different infinities in our calculation of the ground state energy. One is the infrared (IR) divergence, which we can resolve by putting the theory in a box of size L with periodic boundary conditions. By regulating[8] the IR divergence in this way we can

[8] Regulation is the process of modifying our calculation to make the result finite. We will often need to introduce regulators in intermediate steps in our computations. There are many ways to regulate. Perhaps surprisingly, many regulators will yield the same final answer. Regulation will play a larger role near the end of the course when we study renormalization. (Most of the "constants" in quantum field theories secretly depend on the energy at which you measure them. This phenomenon is known as renormalization.)

understand that the $\delta(\mathbf{0})$ divergence comes from infinite volume of space

$$\begin{aligned}(2\pi)^3\delta(\mathbf{0}) &= \lim_{L\to\infty}\int_{-\frac{L}{2}}^{\frac{L}{2}} d^3x e^{i\mathbf{x}\cdot\mathbf{p}}|_{\mathbf{p}=0} \\ &= \lim_{L\to\infty}\int_{-\frac{L}{2}}^{\frac{L}{2}} d^3x \\ &= V\,.\end{aligned} \tag{55}$$

The total energy diverges if the volume diverges unless the energy density vanishes $\rho_0 = 0$ so it is not surprising that the total energy diverges in Minkowski space.

Let us compute the energy density of the vacuum. The energy density of the vacuum still diverges

$$\rho_0 = \frac{E_0}{V} = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2}E_{\mathbf{k}} = \infty\,, \tag{56}$$

because we are summing the zero point energy of infinitely many harmonic oscillators. This is an ultraviolet (UV) divergence (the integral diverges for large momentum $|\mathbf{k}|$) which arises because we assumed our theory was valid to arbitrarily high energies. If the theory breaks down at some high energy scale Λ, this divergence could be avoided. We will pursue this idea further in the last part of this course when we discuss renormalization. For non-gravitational physics, only energy differences are observable so this divergence in the vacuum energy is not a serious problem. To avoid this problem we will normal order our operators.

Normal ordering means putting all the annihilation operators to the right without using the commutation relations[9]. For example

$$:a_{\mathbf{k}}a^\dagger_{\mathbf{q}}: = a^\dagger_{\mathbf{q}}a_{\mathbf{k}}\,. \tag{58}$$

The normal-ordered Hamiltonian is

$$:H: = \int \frac{d^3k}{(2\pi)^3 2E_{\mathbf{k}}}E_{\mathbf{k}}(a^\dagger_{\mathbf{k}}a_{\mathbf{k}})\,. \tag{59}$$

The normal-ordered Hamiltonian annihilates the vacuum state

$$:H:|0\rangle = 0|0\rangle\,. \tag{60}$$

[9] Alternatively normal ordering can be defined by subtracting off the vacuum expectation value. For an operator $\mathcal{O}$, the normal-ordered operator is

$$:\mathcal{O}: = \mathcal{O} - \langle 0|\mathcal{O}|0\rangle\,. \tag{57}$$

The two definitions disagree when the operator in question is a complex number. If we use the definition involving the reordering of the creation and annihilation operators then complex numbers are normal ordered, e.g. $:1: = 1$. If we use the definition involving subtracting off the vacuum expectation value then zero is the normal-ordering of any complex number so we have the odd-looking equation $:1: = 0$. We will not encounter situations in this course where this difference is important.

4.2.2. Casimir Effect

Let us take a small detour to understand the energy of the vacuum state more thoroughly.

The energy of the vacuum is not directly observable. However, a change in the vacuum energy produces a force which can be measured. For simplicity we will work in $1+1$ dimensions. If we have two plates separated by a distance a and the vacuum energy in between the plates depends on a, then there will be a force on the plates. Moving the walls of the box will also change the vacuum energy outside of the plates. Making space periodic with some period L makes it easier to account for the vacuum energy outside of the plates. We consider plates at separated by a distance a with $L \gg a$. In a box of size r, the allowed frequencies are quantized $\omega_n = \frac{n\pi}{r}$. The vacuum energy in between two plates separated by a distance r due to a massless scalar field[10] is

$$E(r) = \sum_{n=1}^{\infty} \frac{1}{2}\omega_n \tag{61}$$

and the total vacuum energy of our two plate setup is

$$E_0 = E(a) + E(L-a)\,. \tag{62}$$

The force is

$$\begin{aligned} F &= -\frac{dE_0}{da} \\ &= \left(\frac{1}{a^2} - \frac{1}{(L-a)^2}\right)\frac{\pi}{2}\sum_{n=1}^{\infty} n \\ &= \infty\,. \end{aligned} \tag{63}$$

We have clearly made a mistake somewhere as the force between the two plates is observable and hence should be finite.

Our mistake is that we neglected to think physically about the quantization of the scalar field modes. The quantization is due to the interaction with the plates and sufficiently high frequency modes with very short wavelengths will not interact appreciably with the plates. A physical plate is made of atoms which have some characteristic distance between them. Wavelengths sufficiently shorter than the spacing between the atoms will pass through the plates. Suppose there is some maximum frequency Λ and only modes with $\omega < \pi\Lambda$ interact with the plates. If the high frequency modes simply pass through the plates, then the vacuum energy becomes

$$\begin{aligned} E(r) &= \sum_{n=1}^{n_{max}} \frac{1}{2}\omega_n \\ &= \frac{\pi}{2r}\frac{n_{\mathrm{max}}(n_{\mathrm{max}}+1)}{2} \end{aligned} \tag{64}$$

[10] Not much in the following discussion would change if we considered the electromagnetic field instead. The final answer would differ by a factor of 2.

where $n_{max} = \lfloor \Lambda r \rfloor$ is the greatest integer less than Λr. After some algebra the total energy becomes

$$E_0(a) = \frac{\pi}{4} L\Lambda^2 - \frac{\pi}{4a} x(1-x) \tag{65}$$

where $x = \Lambda a - \lfloor \Lambda a \rfloor$ is the fractional part of Λa. The energy oscillates as a changes. As Λ increases with a fixed, the number of oscillations increases. In the limit $\Lambda \to \infty$, we can average over the oscillations using $\int x(1-x)dx = \frac{1}{6}$ to get

$$E_0(a) = \frac{\pi}{4} L\Lambda^2 - \frac{\pi}{24a} \,. \tag{66}$$

The force is finite and attractive

$$F(a) = -\frac{\pi}{24a^2} \,. \tag{67}$$

The above way of making the sum $\sum \omega_n$ finite ("regulating" the sum) may seem arbitrary and you may be concerned that a different regularization scheme (or different model for how the high frequency modes interact with the plate) could lead to a different result for the force. As you will show in tutorial, the force is independent of the regulator (for a wide class of regulators).

We can predict the attractive force between two plates without having a detailed model for the short-distance interactions between the plate and the field. This separation of scales is both common and extremely fortunate for the progress of science.

4.2.3. Particles

We can construct other states by acting with creation operators on the vacuum. One-particle states of definite energy $E_{\mathbf{k}}$ and momentum $|\mathbf{k}\rangle$, as we will study in the quiz, are

$$|\mathbf{k}\rangle = a^\dagger_{\mathbf{k}} |0\rangle \,. \tag{68}$$

These one particle states are *not* unit normalized

$$\begin{aligned} \langle \mathbf{p}|\mathbf{k}\rangle &= \langle 0| a_{\mathbf{p}} a^\dagger_{\mathbf{k}} |0\rangle \\ &= (2\pi)^3 2E_{\mathbf{k}} \delta(\mathbf{p}-\mathbf{k}) \,. \end{aligned} \tag{69}$$

We have instead chosen a Lorentz-covariant normalization for the states that leads to Lorentz-invariant amplitudes.

We can construct N-particle states by acting with several creation operators:

$$|\mathbf{k}_1, \ldots, \mathbf{k}_N\rangle = a^\dagger_{\mathbf{k}_1} \cdots a^\dagger_{\mathbf{k}_N} |0\rangle \,. \tag{70}$$

Since the creation operators commute, the particles are bosons. The number operator N

$$N = \int \frac{d^3k}{(2\pi)^3 2E_{\mathbf{k}}} a^\dagger_{\mathbf{k}} a_{\mathbf{k}} \tag{71}$$

counts the number of particles in a state

$$N|\mathbf{k}_1,\ldots,\mathbf{k}_M\rangle = M|\mathbf{k}_1,\ldots,\mathbf{k}_M\rangle\,. \tag{72}$$

Note that the normal-ordered Hamiltonian and the number operator commute

$$[:H:,N] = 0 \tag{73}$$

so particle number is conserved. Particle number is conserved because we have no interactions. Particle number will *not* be conserved in the interacting theories we will study.

The state space is a Fock space or the direct sum of n-particle Hilbert spaces

$$\mathcal{F} = \oplus_{n=0}^{\infty}\mathcal{H}_n \tag{74}$$

where $\mathcal{H}_n$ is the n-particle Hilbert space.

4.3. Heisenberg Picture

So far we have been working in the Schrödinger picture. At times it will be useful to switch to the Heisenberg picture where the operators are time-dependent. Time-dependent operators $\mathcal{O}(t)$ are related to time-independent operators $\mathcal{O}$ in the usual way

$$\mathcal{O}(t) = e^{iHt}\mathcal{O}e^{-iHt}\,. \tag{75}$$

The Heisenberg picture annihilation operator can be computed using the BCH formula

$$e^{A}Be^{-A} = B + [A,B] + \frac{1}{2}[A,[A,B]] + \cdots \tag{76}$$

and

$$[H,a_{\mathbf{p}}] = -E_{\mathbf{p}}a_{\mathbf{p}}\,. \tag{77}$$

The result is

$$\begin{aligned} a_{\mathbf{p}}(t) &= e^{iHt}a_{\mathbf{p}}e^{-iHt} \\ &= e^{-iE_{\mathbf{p}}t}a_{\mathbf{p}}\,. \end{aligned} \tag{78}$$

Similarly the Heisenberg picture creation operator is:

$$a^{\dagger}_{\mathbf{p}}(t) = e^{+iE_{\mathbf{p}}t}a^{\dagger}_{\mathbf{p}}\,. \tag{79}$$

The Heisenberg picture field operator is related to the creation and annihilation operators by

$$\varphi(t,\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3 2E_{\mathbf{k}}}\left[a_{\mathbf{k}}e^{-ik\cdot x} + a^{\dagger}_{\mathbf{k}}e^{+ik\cdot x}\right]. \tag{80}$$

We can take this equation as the definition of φ. The physical content of the theory is in the algebra of the creation and annihilation operators and in the Hamiltonian.

We can check that the field operator obeys the Heisenberg equation of motion

$$\begin{aligned}[:H:,\varphi] &= \int \frac{d^3k}{(2\pi)^3 2E_{\mathbf{k}}} \int \frac{d^3p}{(2\pi)^3 2E_{\mathbf{p}}} \left[E_{\mathbf{k}} a^\dagger_{\mathbf{k}} a_{\mathbf{k}}, a_{\mathbf{p}} e^{-ip\cdot x} + a^\dagger_{\mathbf{p}} e^{+ip\cdot x}\right] \\ &= \int \frac{d^3k}{(2\pi)^3 2E_{\mathbf{k}}} \left(-E_{\mathbf{k}} a_{\mathbf{k}} e^{-ik\cdot x} + E_{\mathbf{k}} a^\dagger_{\mathbf{k}} e^{+ik\cdot x}\right) \\ &= -i\partial_t \varphi(t,\mathbf{x})\,. \end{aligned} \tag{81}$$

An interacting field Φ can be related to the creation and annihilation operators in the same way as a free field

$$\Phi(t,\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3 2E_{\mathbf{k}}} \left[\tilde{a}_{\mathbf{k}}(t) e^{-ik\cdot x} + \tilde{a}^\dagger_{\mathbf{k}}(t) e^{+ik\cdot x}\right]. \tag{82}$$

The creation and annihilation operators $\tilde{a}^\dagger_{\mathbf{p}}(t)$ and $\tilde{a}_{\mathbf{p}}(t)$ satisfy the same algebra as the free theory. At a fixed time t_0 we can define the exact creation and annihilation operators to be equal to the free creation operators:

$$\begin{aligned} \tilde{a}^\dagger_{\mathbf{p}}(t_0) &= a^\dagger_{\mathbf{p}} \\ \tilde{a}_{\mathbf{p}}(t_0) &= a_{\mathbf{p}}\,. \end{aligned} \tag{83}$$

The difference between the free and the interacting theories is all in the time dependence of the operators. In the interacting theory the operators that create momentum eigenstates mix as time evolves.

For the remainder of the course we will focus on interacting theories. Unless indicated otherwise, both ϕ and Φ will indicate interacting fields and we will use subscript zeros to indicate free fields[11] ϕ_0 or Φ_0. Going forward we will drop the tildes and the creation operators for interacting real fields will be denoted $a^\dagger_{\mathbf{p}}(t)$.

5. LSZ Reduction Formula

5.1. Derivation

One of the main quantities of interest in scattering problems is the scattering amplitude $\langle f|S|i\rangle$

$$\langle f|S|i\rangle_{\text{Heisenberg}} = \langle f; t=\infty | i; t=-\infty\rangle_{\text{Schrödinger}}\,. \tag{84}$$

[11] Or more precisely interaction-picture field operators.

In the Heisenberg picture, $|i\rangle$ is the initial state, $|f\rangle$ is the final state, and S is the time-evolution operator (in the limit that that the time interval is very long). In scattering applications, we are most often interested in the case that the initial and final states are momentum eigenstates. Momentum eigenstates can be expressed in terms of the creation operators $a^\dagger_{\mathbf{p}}(t)$. For example a two-particle initial state is

$$|i\rangle = a^\dagger_{\mathbf{p}_1}(0)a^\dagger_{\mathbf{p}_2}(0)|\Omega\rangle \tag{85}$$

where $|\Omega\rangle$ is the vacuum state. Here we assume that the Heisenberg and Schrödinger picture operators agree at $t = 0$.

Our main result in this chapter will be a formula relating scattering amplitudes $\langle f|S|i\rangle$ and time-ordered expectation values or Green's functions $\langle\Omega|\mathrm{T}\varphi(x_1)\cdots\varphi(x_N)|\Omega\rangle$ where $|\Omega\rangle$ is the vacuum of the interacting theory. These Green's functions are the subject of the next chapter.

We will make the assumption that apart from the time dependence of the creation and annihilation operators, the creation, annihilation, and field operators in the interacting theory behave just as they do in the free theory. In particular we assume that $a^\dagger_{\mathbf{p}}(t)$ creates a particle with momentum $\mathbf{p}$ at time t and that the creation and annihilation operators obey the same commutation relation at any fixed time

$$[a_{\mathbf{p}}(t), a^\dagger_{\mathbf{p}'}(t)] = (2\pi)^3 2E_{\mathbf{p}}\delta^{(3)}(\mathbf{p}-\mathbf{p}')\,. \tag{86}$$

We also assume that $\varphi(x)|\Omega\rangle$ is a one particle state as $t \to \pm\infty$ and that the form factor is the same as in the free theory

$$\langle\mathbf{k}|\varphi(x)|\Omega\rangle = e^{ik\cdot x}\,. \tag{87}$$

These assumptions imply that the field operator is linear in the creation operator and fix the coefficient of the creation operator. If the field is real then these assumptions also imply that the field operator is linear in the annihilation operator and fix the coefficient of the annihilation operator. The field operator has to take the same form in the interacting theory as in the free theory

$$\varphi(x) = \int \frac{d^3p}{(2\pi)^3 2E_{\mathbf{p}}}\left(a_{\mathbf{p}}(t)e^{-ip\cdot x} + a^\dagger_{\mathbf{p}}(t)e^{+ip\cdot x}\right)\,. \tag{88}$$

The only difference compared to the free field is the time dependence which can be determined from the Heisenberg equation of motion $i\partial_t\varphi(x) = [\varphi, H]$. If we also assume that the conjugate momentum operator $\pi(x)$ is the same as in free theory in the asymptotic past and future, then as you showed in the quiz

$$\lim_{t\to\pm\infty} a^\dagger_{\mathbf{p}}(t) = -i\int d^3x\left(e^{-ik\cdot x}(\partial_t\varphi(x) + iE_{\mathbf{k}}\varphi(x))\right)\,. \tag{89}$$

We will check the consistency of these assumptions later.

The simplest interesting case of scattering is two-to-two scattering, where two initial particles interact briefly and in the asymptotic future there are two final state particles. We can consider the initial and final states:

$$\begin{aligned} |i\rangle &= a_1^\dagger(0)a_2^\dagger(0)|\Omega\rangle \\ |f\rangle &= a_3^\dagger(0)a_4^\dagger(0)|\Omega\rangle \end{aligned} \tag{90}$$

where $a_i(t) \equiv a_{\mathbf{k}_i}(t)$ and $a_i^\dagger(t) \equiv a_{\mathbf{k}_i}^\dagger(t)$. Our goal is to compute the amplitude

$$\langle f|S|i\rangle = \langle\Omega|a_4(+\infty)a_3(+\infty)a_1^\dagger(-\infty)a_2^\dagger(-\infty)|\Omega\rangle \tag{91}$$

since the time evolution operator evolves the creation operators to the far past and the annihilation operators to the far future.

We will compute this amplitude using the useful identity

$$a_1^\dagger(+\infty) - a_1^\dagger(-\infty) = -i\int d^4x e^{-ik\cdot x}(\partial^2 + m^2)\varphi(x) \equiv I_1^\dagger\,. \tag{92}$$

Note that for an interacting theory the Klein-Gordon operator does not annihilate the field operator $(-\partial^2 + m^2)\varphi(x) \neq 0$.

Let us prove the useful identity

$$\begin{aligned} a_1^\dagger(+\infty) - a_1^\dagger(-\infty) &= \int_{-\infty}^{\infty} dt\partial_0 a_1^\dagger(t) \\ &= -i\int d^4x\partial_0(e^{-ik_1\cdot x}(\partial_0\varphi(x) + iE_{\mathbf{k}_1}\varphi(x))) \\ &= -i\int d^4x e^{-ik_1\cdot x}(\partial_0^2 + E_{\mathbf{k}_1}^2)\varphi(x) \\ &= -i\int d^4x e^{-ik_1\cdot x}(\partial_0^2 - \nabla^2 + m^2)\varphi(x) \\ &= -i\int d^4x e^{-ik_1\cdot x}(\partial^2 + m^2)\varphi(x)\,. \end{aligned} \tag{93}$$

In the third line the other two terms coming from applying the product rule of the partial derivative cancel each other. In the fourth line we integrated by parts to get the ∇^2 term.

From this derivation we see that the mass m in the useful identity (92) is the physical mass of the particle. The mass entered when we used the relativistic dispersion relation $E_{\mathbf{k}}^2 = m^2 + \mathbf{k}^2$. At this point we have not specified a Lagrangian. In chapter 8 we will see that the mass parameter in the Lagrangian is not necessarily equal to the physical mass.

Similarly for the annihilation operator we have

$$a_1(+\infty) - a_1(-\infty) = i\int d^4x e^{+ik_1\cdot x}(\partial^2 + m^2)\varphi(x) \equiv I_1\,. \tag{94}$$

Now we can rewrite the scattering amplitude as

$$\begin{aligned}\langle f|S|i\rangle &= \langle\Omega|a_4(+\infty)a_3(+\infty)a_1^\dagger(-\infty)a_2^\dagger(-\infty)|\Omega\rangle \\ &= \langle\Omega|\mathrm{T}a_4(+\infty)a_3(+\infty)a_1^\dagger(-\infty)a_2^\dagger(-\infty)|\Omega\rangle\end{aligned} \tag{95}$$

where the time ordering symbol T indicates the later operator is to the left. For example for two field operators

$$T\varphi(x)\varphi(y) = \begin{cases}\varphi(x)\varphi(y) & x^0 > y^0 \\ \varphi(y)\varphi(x) & y^0 > x^0\,.\end{cases} \tag{96}$$

Now using the identities we derived above

$$\begin{aligned}a_1^\dagger(-\infty) &= a_1^\dagger(+\infty) - I_1^\dagger \\ a_4(+\infty) &= a_4(-\infty) + I_4\,,\end{aligned} \tag{97}$$

we have

$$\begin{aligned}\langle f|S|i\rangle &= \langle\Omega|\mathrm{T}a_4(-\infty)a_3(-\infty)a_1^\dagger(+\infty)a_2^\dagger(+\infty)|\Omega\rangle \\ &\quad + \ldots + \langle\Omega|\mathrm{T}I_4I_3I_1^\dagger I_2^\dagger|\Omega\rangle\,.\end{aligned} \tag{98}$$

The time ordering brings the annihilation operators to the right, so all terms except the last vanish. This leads to the LSZ reduction formula[12]:

$$\langle f|S|i\rangle = i^{2+2}\int\left(\prod_{j=1}^{2+2}d^4x_j e^{-i\lambda_j k_j\cdot x_j}(\partial_j^2+m^2)\right)\langle\Omega|\mathrm{T}\varphi_1\varphi_2\varphi_3\varphi_4|\Omega\rangle \tag{99}$$

where λ_j is 1 for initial states and -1 for final states. We introduced the notation $\partial_j^2 \equiv \eta^{\mu\nu}\frac{\partial}{\partial x_j^\mu}\frac{\partial}{\partial x_j^\nu}$ and $\varphi_1 \equiv \varphi(x_1)$. This works for $n \to m$ scattering with suitable modifications.

5.2. Checking Assumptions

We assumed $a_1^\dagger(\pm\infty)|\Omega\rangle$ and $\lim_{t\to\pm\infty}\varphi(x)|\Omega\rangle$ are one particle states. These assumptions justified our use of the equation (88) relating the field operators and creation and annihilation operators.

We can define the creation and annihilation operators to be the operators that add or subtract a particle at a given time, but it is certainly questionable whether $\varphi(x)|\Omega\rangle$ is a one-particle state in the asymptotic past or future. Indeed particles can decay so one- and multi-particle states should not be orthogonal. Nevertheless we should have some way to be able to differentiate between one- and multi-particle states. If we want $\lim_{t\to\pm\infty}\varphi(x)|\Omega\rangle$ to be a one-particle state with the correct form factor then we need to check that:

[12]Switching the time-ordering and derivatives is not strictly justified, but the end result is correct.

- The one-particle state $\varphi(x)|\Omega\rangle$ as $t \to \pm\infty$ is orthogonal to the vacuum state $\langle\Omega|\varphi(x)|\Omega\rangle = 0$.
- The form factor is correct $\langle \mathbf{k}|\varphi(x)|\Omega\rangle = e^{ik\cdot x}$.
- We can distinguish one-particle and multi-particle states at asymptotic times.

5.2.1. $\langle\Omega|\varphi(x)|\Omega\rangle = 0$

The overlap between the one-particle state $\lim_{t\to\pm\infty}\varphi(x)|\Omega\rangle$ and the zero-particle vacuum $|\Omega\rangle$ should be zero. To compute the overlap we can write the field operator using the spacetime translation operator

$$\varphi(x) = e^{iP\cdot x}\varphi(0)e^{-iP\cdot x}\,. \tag{100}$$

The overlap becomes

$$\begin{aligned}\langle\Omega|\varphi(x)|\Omega\rangle &= \langle\Omega|e^{iP\cdot x}\varphi(0)e^{-iP\cdot x}|\Omega\rangle \\ &= \langle\Omega|\varphi(0)|\Omega\rangle\,.\end{aligned} \tag{101}$$

The last line is due to the translational invariance of $|\Omega\rangle$. There is no reason for $\langle\Omega|\varphi(0)|\Omega\rangle$ to vanish in general. It is a Lorentz invariant number so

$$\langle\Omega|\varphi(x)|\Omega\rangle = v\,. \tag{102}$$

All is not lost. If we want to use the LSZ reduction formula, we can redefine $\tilde{\varphi}(x) = \varphi(x) - v$ so

$$\langle\Omega|\tilde{\varphi}(x)|\Omega\rangle = 0\,. \tag{103}$$

We will drop the tilde from now on.

5.2.2. $\langle \mathbf{k}|\varphi(x)|\Omega\rangle = e^{ik\cdot x}$

We can compute the form factor in the same way

$$\begin{aligned}\langle \mathbf{k}|\varphi(x)|\Omega\rangle &= \langle \mathbf{k}|e^{iP\cdot x}\varphi(0)e^{-iP\cdot x}|\Omega\rangle \\ &= e^{ik\cdot x}\langle \mathbf{k}|\varphi(0)|\Omega\rangle\,.\end{aligned} \tag{104}$$

Again we used that the vacuum state is translationally invariant. Since our one-particle states are Lorentz covariant, we expect $\langle \mathbf{k}|\varphi(0)|\Omega\rangle$ to be Lorentz invariant, and thus a function of $k^2 = m^2$. If we want our one-particle states to be correctly normalized, we want

$$\langle \mathbf{k}|\varphi(0)|\Omega\rangle = 1\,. \tag{105}$$

Again there is no reason that $\langle \mathbf{k}|\varphi(0)|\Omega\rangle$ has to be unity. We can fix this problem by rescaling $\tilde{\varphi}(x) = N\varphi(x)$, where N is a constant. We will drop the tilde again.

5.2.3. Single- Versus Multi-Particle States

We can label a multiparticle state by the total 3-momentum $\mathbf{p}$ and some other parameters which we will collectively denote by σ. Let us calculate the overlap with $\varphi(x)|\Omega\rangle$

$$\begin{aligned}\langle \mathbf{p},\sigma|\varphi(x)|\Omega\rangle &= \langle \mathbf{p},\sigma|e^{iP\cdot x}\varphi(0)e^{-iP\cdot x}|\Omega\rangle \\ &= e^{ip\cdot x}\langle \mathbf{p},\sigma|\varphi(0)|\Omega\rangle \\ &= e^{ip\cdot x}A_\sigma(\mathbf{p})\,.\end{aligned} \tag{106}$$

In general this is non-zero! There is no need to panic yet. We want to be able to describe the decay of a single particle, which would be difficult if this overlap always vanished. Demanding that (106) vanishes is too strong of a condition. We really only need

$$\langle \mathbf{p},\sigma|a_1^\dagger(\pm\infty)|\Omega\rangle = 0 \tag{107}$$

so that our creation operator does not create multiparticle states at asymptotic times. Even this condition is a slightly more than we need. Everything will be fine if

$$\langle \psi|A_1^\dagger(\pm\infty)|\Omega\rangle = 0 \tag{108}$$

where $|\psi\rangle$ is a normalizable state

$$|\psi\rangle = \sum\!\!\!\!\!\!\int_\sigma d^3p\psi_\sigma(\mathbf{p})|\mathbf{p},\sigma\rangle \tag{109}$$

and

$$A_1^\dagger(t)|\Omega\rangle = \int d^3p f_1(\mathbf{p})a_1^\dagger(t)|\Omega\rangle \tag{110}$$

where $\psi_\sigma(\mathbf{p})$ and $f_1(\mathbf{p})$ are wavepackets. In general σ contains some discrete parameters such as particle number and some continuous parameters such as relative momenta, and so we are both summing and integrating over σ. The wavepackets could be gaussians (e.g. $f_1(\mathbf{p}) \propto \exp[-(\mathbf{k}-\mathbf{k}_1)^2/\Sigma^2]$) but their precise form will be unimportant. With this definition

$$\begin{aligned}\langle \psi|A_1^\dagger(t)|\Omega\rangle = -i\sum\!\!\!\!\!\!\int_\sigma d^3p\psi_\sigma^*(\mathbf{p})\int d^3k f_1(\mathbf{k})\int d^3x \\ (e^{-ik\cdot x}\partial_0\langle \mathbf{p},\sigma|\varphi(x)|\Omega\rangle - \langle \mathbf{p},\sigma|\varphi(x)|\Omega\rangle\partial_0 e^{-ik\cdot x})\,.\end{aligned} \tag{111}$$

If we use the result (106) then the integrand of the right hand side of (111) becomes

$$i(p^0+k^0)e^{i(p-k)\cdot x}A_\sigma(\mathbf{p})\,. \tag{112}$$

Now we use

$$\int d^3x e^{i(\mathbf{k}-\mathbf{p})\cdot\mathbf{x}} = (2\pi)^3\delta^{(3)}(\mathbf{k}-\mathbf{p}) \tag{113}$$

to perform the integral over $\mathbf{k}$. We obtain

$$\langle\psi|A_1^\dagger(t)|\Omega\rangle = \sum_\sigma \int d^3p(2\pi)^3(p^0+k^0)\psi_\sigma^*(\mathbf{p})f_1(\mathbf{p})A_\sigma(\mathbf{p})e^{i(p^0-k^0)t} \tag{114}$$

where

$$\begin{aligned} p^0 &= (\mathbf{p}^2+M^2)^{1/2} \\ k^0 &= (\mathbf{p}^2+m^2)^{1/2}\,. \end{aligned} \tag{115}$$

Here M is the invariant mass of n-particle state. If there are no bound states then

$$M \geq 2m > m \tag{116}$$

so $p^0 > k^0$. The exponential factor $e^{i(p^0-k^0)t}$ oscillates more and more rapidly as $t \to \pm\infty$, so the $\mathbf{p}$ integral vanishes as $t \to \pm\infty$ [13]. We end up with the desired result

$$\langle\psi|A_1^\dagger(\pm\infty)|\Omega\rangle = 0\,. \tag{117}$$

Physically this equation says that one particle wavepackets and multiparticle wavepackets spread out differently and the overlap between the wavepackets goes to zero as $t \to \pm\infty$. In this sense single- and multi-particle states can be distinguished.

5.2.4. Discussion

Suppose our original theory is

$$\mathcal{L} = \frac{1}{2}(\partial\varphi)^2 - \frac{1}{2}m^2\varphi^2 + \frac{1}{3!}g\varphi^3 \tag{118}$$

We have to shift, rescale, and rename the field to use the LSZ formula:

$$\mathcal{L} = \frac{1}{2}Z_\varphi(\partial\varphi)^2 - \frac{1}{2}Z_m m^2\varphi^2 + \frac{1}{3!}Z_g g\varphi^3 + Y\varphi \tag{119}$$

Note that a new term $Y\varphi$ appears. Renormalization is required! We will return to this issue in chapter 8. We will ignore this issue in chapters 6 and 7 and assume that

$$\begin{aligned} \langle\Omega|\varphi(x)|\Omega\rangle &= 0 \\ \langle\mathbf{k}|\varphi(x)|\Omega\rangle &= e^{ik\cdot x}\,. \end{aligned} \tag{120}$$

We will also assume all $Z_i = 1$ and $Y_j = 0$ in chapters 6 and 7. In general these are not consistent assumptions beyond leading order in perturbation theory. It is consistent to make these assumptions at leading order in perturbation theory (at tree level).

In the last chapter we related cross sections and decay rates to scattering amplitudes. We saw that in order to compute $\langle f|S|i\rangle$ we need to compute $\langle\Omega|\mathrm{T}\varphi_1\ldots\varphi_n|\Omega\rangle$. Next we will learn how to compute vacuum expectation values of time-ordered products of fields.

[13] due to the Riemann-Lebesgue lemma

6. Time-Ordered Expectation Values

In the previous chapter we saw that if we want to compute a scattering amplitude $\langle f|S|i\rangle$ then we have to compute the time-ordered expectation value or Green's function:

$$\langle\Omega|T\varphi_1\dots\varphi_n|\Omega\rangle \tag{121}$$

where $\varphi_i \equiv \varphi(x_i)$. In this chapter we will compute this time-ordered expectation value perturbatively. We will show that the Green's function in the interacting theory is related to the Green's function in a free theory using the interaction picture. Green's functions in a free theory can be related to the two-point function in the free theory using Wick's theorem. Finally we will compute the two-point function in the free theory (also known as the Feynman propagator).

6.1. Relating Free and Interacting Time-Ordered Expectation Values

In the LSZ formula we encountered the expectation value of time ordered products of Heisenberg picture fields. To evaluate these expectation values it will be useful to go to the interaction picture. We assume the total Hamiltonian H can be expressed as the sum of a free Hamiltonian H_0 plus an interaction Hamiltonian H_1

$$H = H_0 + H_1\,. \tag{122}$$

In this course we will mostly be interested in the case that the free Hamiltonian is the Klein-Gordon Hamiltonian.

The interaction picture field $\varphi_0(\mathbf{x},t)$ is related to the Heisenberg picture field $\varphi(\mathbf{x},t)$ by

$$\begin{aligned}\varphi_0(\mathbf{x},t) &= e^{iH_0(t-t_0)}\varphi(\mathbf{x},t_0)e^{-iH_0(t-t_0)} \\ &= U(t,t_0)\varphi(\mathbf{x},t)U(t,t_0)^\dagger\end{aligned} \tag{123}$$

where $U(t_1,t_2)$ is the interaction picture time evolution operator:

$$U(t_1,t_2) \equiv e^{iH_0(t_1-t_2)}e^{-iH(t_1-t_2)} \tag{124}$$

$$= T\exp\left[-i\int_{t_1}^{t_2} dt H_{1I}(t)\right]. \tag{125}$$

The subscript 0 on the interaction picture fields indicates that the interaction picture fields evolve with the free Hamiltonian and thus behave like free fields. We will often refer to them as free fields in a slight abuse of terminology.

We can use Dyson's formula to compute the Green's functions on the right-hand side of the LSZ reduction formula. To accomplish this goal we also need to relate the free and interacting vacua $|0\rangle$ and $|\Omega\rangle$.

In deriving the LSZ formula we assumed that the annihilation operators annihilated the vacuum at $t = \pm\infty$. In the Schrödinger picture we have

$$\begin{aligned} a_{\mathbf{p}}(t_0)|\Omega(t)\rangle\big|_{t=\pm\infty} &= 0 \\ a_{\mathbf{p}}(t_0)e^{-iH(t-t_0)}|\Omega\rangle\Big|_{t=\pm\infty} &= 0\,. \end{aligned} \tag{126}$$

The free annihilation operator annihilates the free vacuum at any time. In particular the free annihilation operator annihilates the free vacuum at $t = \pm\infty$ in the free theory

$$\begin{aligned} a_{\mathbf{p}}(t_0)|0(t)\rangle\big|_{t=\pm\infty} &= 0 \\ a_{\mathbf{p}}(t_0)e^{-iH_0(t-t_0)}|0\rangle\Big|_{t=\pm\infty} &= 0\,. \end{aligned} \tag{127}$$

At time t_0 the free and interacting theory annihilation operators are the same. If the vacuum is unique then the free vacuum and interacting vacuum must be proportional so

$$\begin{aligned} |\Omega\rangle &= N_i \lim_{t\to-\infty} e^{iH(t-t_0)}e^{-iH_0(t-t_0)}|0\rangle \\ &= N_i U_{0-\infty}|0\rangle\,. \end{aligned} \tag{128}$$

where N_i is a normalization constant. Similarly

$$\langle\Omega| = N_f\langle 0|U_{\infty 0}\,. \tag{129}$$

The normalization constants N_i and N_f are in general different. Combining our results (assuming $t_1 > t_2 > \ldots > t_n$ without loss of generality) we can write the Heisenberg picture Green's function in the interaction picture

$$\begin{aligned} \langle\Omega|\mathrm{T}\varphi(x_1)\varphi(x_2)\ldots\varphi(x_n)|\Omega\rangle &= \langle\Omega|\varphi(x_1)\varphi(x_2)\ldots\varphi(x_n)|\Omega\rangle \\ &= N_iN_f\langle 0|U_{\infty 0}U_{01}\varphi_0(x_1)U_{10}U_{02}\ldots U_{0n}\varphi_0(x_n)U_{n0}U_{0-\infty}|0\rangle \\ &= N_iN_f\langle 0|U_{\infty 1}\varphi_0(x_1)U_{12}\varphi_0(x_2)\ldots U_{(n-1)n}\varphi_0(x_n)U_{n-\infty}|0\rangle \\ &= N_iN_f\langle 0|\mathrm{T}U_{\infty 1}\varphi_0(x_1)U_{12}\varphi_0(x_2)\ldots U_{(n-1)n}\varphi_0(x_n)U_{n-\infty}|0\rangle \\ &= N_iN_f\langle 0|\mathrm{T}\varphi_0(x_1)\ldots\varphi_0(x_n)U_{\infty-\infty}|0\rangle\,. \end{aligned} \tag{130}$$

In the third line we used the composition property of U. In the fourth line we noted that all of the operators were already time-ordered. In the fifth line we used the composition property again.

We want the vacuum state to be unit normalized $\langle\Omega|\Omega\rangle = 1$. If we compute the amplitude

$$\langle\Omega|\Omega\rangle = N_iN_f\langle 0|U_{\infty-\infty}|0\rangle \tag{131}$$

we can determine the normalization constants

$$N_iN_f = \frac{1}{\langle 0|U_{\infty-\infty}|0\rangle}\,. \tag{132}$$

Putting our results so far together we have

$$\langle\Omega|\mathrm{T}\varphi(x_1)\varphi(x_2)\dots\varphi(x_n)|\Omega\rangle = \frac{\langle 0|\mathrm{T}\varphi_0(x_1)\dots\varphi_0(x_n)e^{-i\int_{-\infty}^{\infty}dtH_{1I}(t)}|0\rangle}{\langle 0|\mathrm{T}e^{-i\int_{-\infty}^{\infty}dtH_{1I}(t)}|0\rangle}\,. \tag{133}$$

Our next task is to write the interaction Hamiltonian in terms of free fields. Suppose

$$H_1(t) = \int d^3x\frac{\lambda}{4!}\varphi^4(\mathbf{x},t) \tag{134}$$

then

$$\begin{aligned} H_1(t_0) &= \int d^3x\frac{\lambda}{4!}\varphi^4(\mathbf{x},t_0) \\ &= \int d^3x\frac{\lambda}{4!}\varphi_0^4(\mathbf{x},t_0)\,. \end{aligned} \tag{135}$$

Using the time evolution of the interaction picture fields

$$\varphi_0(\mathbf{x},t) = e^{iH_0(t-t_0)}\varphi_0(\mathbf{x},t_0)e^{-iH_0(t-t_0)} \tag{136}$$

we can write the interaction picture interaction Hamiltonian

$$\begin{aligned} H_{1I}(t) &= e^{iH_0(t-t_0)}\int d^3x\frac{\lambda}{4!}\varphi_0^4(\mathbf{x},t_0)e^{-iH_0(t-t_0)} \\ &= \int d^3x\frac{\lambda}{4!}\varphi_0^4(\mathbf{x},t)\,. \end{aligned} \tag{137}$$

The interaction picture interaction Hamiltonian takes the same form as the Heisenberg picture interaction Hamiltonian, with interacting fields replaced with free fields.

It is usually more convenient to write the time evolution operator in terms of the manifestly Lorentz-invariant Lagrangian instead of the Hamiltonian

$$\begin{aligned} U_{\infty,-\infty} &= \mathrm{T}\exp\left[-i\int_{-\infty}^{\infty}dtH_{1I}(t)\right] \\ &= \mathrm{T}\exp\left[i\int_{-\infty}^{\infty}d^4x\mathcal{L}_{\text{int}}[\varphi_0]\right] \end{aligned} \tag{138}$$

where $\mathcal{L}_{\text{int}} = \mathcal{L} - \mathcal{L}_0$.

Combining our results we find that Dyson's formula lets us relate interacting Green's functions to free Green's functions

$$\langle\Omega|\mathrm{T}\varphi(x_1)\cdots\varphi(x_n)|\Omega\rangle = \frac{\langle 0|\mathrm{T}\varphi_0(x_1)\cdots\varphi_0(x_n)\exp\left[i\int d^4x\mathcal{L}_{\text{int}}[\varphi_0]\right]|0\rangle}{\langle 0|\mathrm{T}\exp\left[i\int d^4x\mathcal{L}_{\text{int}}[\varphi_0]\right]|0\rangle}\,. \tag{139}$$

This is an exact expression. If we can compute the right-hand side of this equation (139) then we can compute scattering amplitudes using the LSZ reduction formula. Computing

the right-hand side of (139) is an easier task than computing the left-hand side directly because interaction picture fields behave like free fields.

Unfortunately we usually cannot compute the right-hand side exactly. However we can compute it perturbatively by expanding the exponential

$$\exp\left[i\int d^4x\mathcal{L}_{\text{int}}[\varphi_0]\right] = 1 + i\int d^4x\mathcal{L}_{\text{int}}[\varphi_0] + \frac{i^2}{2}\left(\int d^4x\mathcal{L}_{\text{int}}[\varphi_0]\right)^2 + \cdots. \tag{140}$$

Using this expansion of the exponential, we can find a series expansion for the right-hand side of (139). Each term in this series expansion will involve Green's function's in the free theory. We have essentially reduced the problem of computing Green's functions in an interacting theory to computing Green's functions in a free theory.

Now the question is how to evaluate $\langle 0|\mathrm{T}\varphi_0(x_1)\ldots\varphi_0(x_n)|0\rangle$. Wick's Theorem relates the n-point correlation function of free fields to the two-point correlation function of free fields.

For the remainder of the course we will work in the interaction picture unless otherwise noted. To simplify notation we will usually drop the subscript 0's from now on.

6.2. Wick's Theorem

Wick's theorem will allow us to compute $\langle 0|\mathrm{T}\varphi_0(x_1)\ldots\varphi_0(x_n)|0\rangle$. Our strategy will be to turn the time-ordered product of fields into a normal-ordered product of fields. Normal-ordered products have all creation operators to the left, so a vacuum expectation value of a (nontrivial) normal-ordered product vanishes.[14]

It will be useful to separate[15] the field into a part involving the creation operator and a part involving the annihilation operator

$$\varphi(x) = \varphi_+(x) + \varphi_-(x) \tag{141}$$

where annihilation operator part is given by

$$\varphi_-(x) = \int \frac{d^3p}{(2\pi)^3 2E_{\mathbf{p}}} a_{\mathbf{p}} e^{-ip\cdot x} \tag{142}$$

and the creation operator part is given by

$$\varphi_+(x) = \int \frac{d^3p}{(2\pi)^3 2E_{\mathbf{p}}} a^\dagger_{\mathbf{p}} e^{+ip\cdot x}\,. \tag{143}$$

[14] Recall that we can also define a normal-ordered product of operators to be the product of the operators minus the vacuum expectation value of that product of operators.

[15] Historically this was usually done by separating into positive φ^+ and negative φ^- frequency parts. Some of the recommended texts use this notation which is related to the notation used here by $\varphi^+ = \varphi_-$ and $\varphi^- = \varphi_+$.

The simplest nontrivial case is the time-ordered product of two fields $\mathrm{T}\varphi(x)\varphi(y)$. For $x^0 > y^0$ the time-ordered product can be related to the normal-ordered product by

$$\begin{aligned}\mathrm{T}\varphi(x)\varphi(y) &= \varphi(x)\varphi(y) \\ &= (\varphi_-(x) + \varphi_+(x))(\varphi_-(y) + \varphi_+(y)) \\ &= \varphi_-(x)\varphi_-(y) + \varphi_+(y)\varphi_-(x) + [\varphi_-(x), \varphi_+(y)] + \varphi_+(x)\varphi_-(y) + \varphi_+(x)\varphi_+(y)\,.\end{aligned} \tag{144}$$

For $y^0 > x^0$, we get the same final expression except the commutator is $[\varphi_-(y), \varphi_+(x)]$ instead of $[\varphi_-(x), \varphi_+(y)]$. We can combine the two cases if we define the contraction of two fields

$$\overbrace{\varphi(x)\varphi}(y) = \underbrace{\varphi(x)\varphi}(y) \equiv \Theta(x^0 - y^0)[\varphi_-(x), \varphi_+(y)] + \Theta(y^0 - x^0)[\varphi_-(y), \varphi_+(x)]\,. \tag{145}$$

Then we see that the time-ordered product of two fields is the sum of the normal-ordered product of two fields and the contraction of two fields

$$\mathrm{T}\varphi(x)\varphi(y) =: \varphi(x)\varphi(y) : + \overbrace{\varphi(x)\varphi}(y)\,. \tag{146}$$

Since the vacuum expectation value of normal-ordered operators vanishes, the vacuum expectation value of the time-ordered product of two fields (also known as the Feynman propagator Δ_F) is equal to the vacuum expectation value of the contraction of two fields

$$\langle 0|\mathrm{T}\varphi(x)\varphi(y)|0\rangle = \langle 0|\overbrace{\varphi(x)\varphi}(y)|0\rangle = \Delta_F(x - y)\,. \tag{147}$$

In the free theory $[\varphi(x), \varphi(y)]$ is a complex number[16] so $[\varphi(x), \varphi(y)] = \langle 0|[\varphi(x), \varphi(y)]|0\rangle$ and the second equality in (147) holds without taking the vacuum expectation value

$$\overbrace{\varphi(x)\varphi}(y) = \Delta_F(x - y)\,. \tag{148}$$

Wick's theorem generalizes equation (146) to any number of operators

$$\mathrm{T}\varphi_1 \ldots \varphi_n = \sum : \text{all possible contractions} : \,. \tag{149}$$

where $\varphi_i = \varphi(x_i)$. For example

$$\mathrm{T}\varphi_1\varphi_2\varphi_3 =: \varphi_1\varphi_2\varphi_3 : + : \underbrace{\varphi_1}\varphi_2\varphi_3 : + : \underbrace{\varphi_1\varphi_2\varphi_3} : + : \varphi_1\underbrace{\varphi_2}\varphi_3 : \,. \tag{150}$$

Note that :all possible contractions: includes terms that are partially contracted.

[16] To see that the commutator is a complex number, note that the free field operators are linear combinations of the creation and annihilation operators. The commutator of two creation or annihilation operators is a complex number, so the commutator of two free field operators is a complex number.

Sketch of Proof: We showed that Wick's theorem holds for $n = 2$. Assume it works for $n = m - 1$ and without loss of generality take $x_1^0 \geq x_2^0 \geq \cdots \geq x_m^0$ (we can always relabel if this is not the case). Applying Wick's theorem to $\varphi_2 \ldots \varphi_m$ we have

$$\begin{aligned} \mathrm{T}\varphi_1 \ldots \varphi_m &= \varphi_1(\mathrm{T}\varphi_2 \ldots \varphi_m) \\ &= \varphi_1 \sum : \text{contractions not involving } \varphi_1 : \\ &= (\varphi_{1-} + \varphi_{1+}) \sum : \text{contractions not involving } \varphi_1 : . \end{aligned} \tag{151}$$

We can bring φ_{1+} inside the normal-ordering symbols since it contains only creation operators. Look at the first term involving φ_{1-}

$$\begin{aligned} \varphi_{1-} : \varphi_2 \ldots \varphi_m : &=: \varphi_2 \ldots \varphi_m : \varphi_{1-} + [\varphi_{1-}, : \varphi_2 \ldots \varphi_m :] \\ &=: \varphi_{1-}\varphi_2 \ldots \varphi_m : + : [\varphi_{1-}, \varphi_{2+}]\varphi_3 \ldots \varphi_m : + \ldots + : \varphi_2 \ldots [\varphi_{1-}, \varphi_{m+}] : \\ &=: \varphi_{1-}\varphi_2 \ldots \varphi_m : + : \overbrace{\varphi_1\varphi_2} \ldots \varphi_m : + \ldots + : \overbrace{\varphi_1 \ldots \varphi_m} : . \end{aligned} \tag{152}$$

The other terms work similarly.

We know how to express $\langle 0|\mathrm{T}\varphi_1 \ldots \varphi_n|0\rangle$ as a sum of products of Feynman propagators now. Only completely contracted terms survive since

$$\langle 0| : \varphi_1 \ldots \varphi_n : |0\rangle = 0 . \tag{153}$$

For example

$$\langle 0|\mathrm{T}\varphi_1\varphi_2\varphi_3\varphi_4|0\rangle = \Delta_{12}\Delta_{34} + \Delta_{13}\Delta_{24} + \Delta_{14}\Delta_{23} \tag{154}$$

where

$$\Delta_{12} = \Delta_F(x_1 - x_2) . \tag{155}$$

6.3. Propagators

In tutorial 5 you showed that the amplitude for a particle to propagate from y to x is given by

$$\begin{aligned} \langle 0|\varphi(x)\varphi(y)|0\rangle &= \langle 0|[\varphi_-(x), \varphi_+(y)]|0\rangle \\ &= \int \frac{d^3p}{(2\pi)^3 2E_{\mathbf{p}}} \int \frac{d^3k}{(2\pi)^3 2E_{\mathbf{k}}} e^{-ip\cdot x + ik\cdot y} \langle 0|a_{\mathbf{p}} a_{\mathbf{k}}^\dagger|0\rangle \\ &= \int \frac{d^3p}{(2\pi)^3 2E_{\mathbf{p}}} e^{-ip\cdot(x-y)} \end{aligned} \tag{156}$$

where we used that $\langle 0|a_{\mathbf{p}} a_{\mathbf{k}}^\dagger|0\rangle \propto \delta(\mathbf{p} - \mathbf{k})$.

The most useful way to write the Feynman propagator is as a contour integral. The Feynman propagator can be written

$$\begin{aligned}\Delta_F(x-y) &= \langle 0|\mathrm{T}\varphi(x)\varphi(y)|0\rangle \\ &= \int \frac{d^3p}{(2\pi)^3}\left(\frac{1}{2E_{\mathbf{p}}}e^{-ip\cdot(x-y)}|_{p^0=E_{\mathbf{p}}}\Theta(x^0-y^0)+\frac{1}{2E_{\mathbf{p}}}e^{-ip\cdot(x-y)}|_{p^0=-E_{\mathbf{p}}}\Theta(y^0-x^0)\right) \\ &= \int \frac{d^3p}{(2\pi)^3}\int_{C_F}\frac{dp^0}{2\pi i}\frac{-1}{p^2-m^2}e^{-ip\cdot(x-y)}\end{aligned} \tag{157}$$

where the contour C_F runs along the real p^0 axis except for an infinitesimal detours into the lower half plane to avoid the pole at $-E_{\mathbf{p}}$ and an infinitesimal detour into the upper half plane to avoid the pole at $E_{\mathbf{p}}$. This derivation is easiest to understand from the bottom up. To go from the third line to the second line we can complete the contour in the lower half plane for $x^0 > y^0$ or in the upper half plane for $y^0 > x^0$ and use contour integration. To go from the second line to the first line we can integrate over $-\mathbf{p}$ instead of $\mathbf{p}$ for the second term.

We can also write the Feynman propagator as

$$\Delta_F(x-y) = \int \frac{d^4p}{(2\pi)^4}\frac{i}{p^2-m^2+i\epsilon}e^{-ip\cdot(x-y)} \tag{158}$$

where the $+i\epsilon$ is a mnemonic for the contour C_F that goes below the pole at $-E_{\mathbf{p}}$ and above the pole at $E_{\mathbf{p}}$.

The Feynman propagator is a Green's function for the Klein-Gordon operator

$$\begin{aligned}(\partial^2+m^2)\Delta_F(x-y) &= \int_{C_F}\frac{d^4p}{(2\pi)^4}\frac{i}{p^2-m^2}(-p^2+m^2)e^{-ip\cdot(x-y)} \\ &= -i\int_{C_F}\frac{d^4p}{(2\pi)^4}e^{-ip\cdot(x-y)} \\ &= -i\delta(x-y)\end{aligned} \tag{159}$$

The choice of contour is not used in the above derivation. Other choices of contours lead to the more familiar retarded or advanced propagators.

We now know how to compute S-matrix elements perturbatively. In the next chapter we will develop another faster-in-practice but equivalent way to compute S-matrix elements.

7. Feynman Diagrams

Feynman diagrams are a way to graphically represent perturbative calculations[17] in quantum field theory. Each Feynman diagram is associated with an analytic expression. The advantage of Feynman diagrams is their relative simplicity compared to the approach of using Wick's theorem, Dyson's formula, and the LSZ reduction formula directly.

Feynman diagrams can be used to calculate several quantities of interest including Green's functions and matrix elements. The Feynman rules for which diagrams to draw and which analytic expression to associate with each diagram depends on which quantity we want to compute. In all cases we represent spacetime points as dots and propagators as lines.

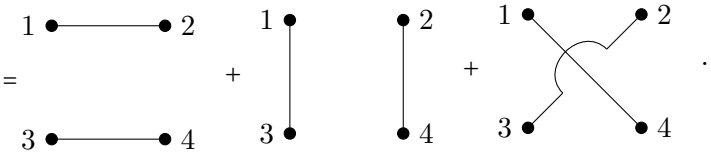
7.1. Position Space Feynman Rules for Green's Functions

In this subsection we will develop Feynman rules in position space for Green's functions. Our strategy will be to determine what the Feynman rules must be to match the results of our analytic calculations for a few examples.

One of the simplest possible Green's functions is the free four-point function. Using Wick's theorem we determined the free four-point function in (154) to be

$$\langle 0|\mathrm{T}\varphi_1\varphi_2\varphi_3\varphi_4|0\rangle = \Delta_{12}\Delta_{34} + \Delta_{13}\Delta_{24} + \Delta_{14}\Delta_{23}\,. \tag{160}$$

By representing each propagator with a line and each spacetime point as a dot this result can be graphically depicted as the sum of three Feynman diagrams

$\langle 0|\mathrm{T}\varphi_1\varphi_2\varphi_3\varphi_4|0\rangle =$  (161)

Note that there is no interaction where the lines cross in the third diagram since there is no dot at the intersection. These diagrams are our first examples of Feynman diagrams.

Before we present the rules from going from the diagram to the expression, let's look another example of a correlation function and see what kind of diagram results.

One of the key results from the previous chapter is how to write the Heisenberg picture Green's function in the interaction picture

$$\langle\Omega|\mathrm{T}\varphi_1\varphi_2|\Omega\rangle = \frac{\langle 0|\mathrm{T}\varphi_1\varphi_2 e^{i\int d^4y\mathcal{L}_{int}}|0\rangle}{\langle 0|\mathrm{T}e^{i\int d^4y\mathcal{L}_{int}}|0\rangle} \equiv G(x_1,x_2)\,. \tag{162}$$

[17] Perturbative calculations in quantum field theory typically yield divergent asymptotic series. These series usually provide a good approximation when truncated after a few terms. Caution is necessary when physically interpreting Feynman diagrams because the diagrams represent terms in the asymptotic series. The sum of all Feynman diagrams typically diverges instead of equaling the quantity of interest.

For the purpose of illustration we will look at φ^4 theory with an interaction term

$$\mathcal{L}_{\rm int}(y) = \frac{-\lambda}{4!}\varphi^4(y)\,. \tag{163}$$

The first two terms in the numerator of $G(x_1, x_2)$ are

$$\langle 0|\mathrm{T}\varphi_1\varphi_2 + \mathrm{T}\varphi_1\varphi_2\frac{(-i\lambda)}{4!}\int d^4y\varphi^4(y)|0\rangle\,. \tag{164}$$

The first term can be represented by the diagram

$$\langle 0|\mathrm{T}\varphi_1\varphi_2|0\rangle = \quad 1 \bullet\!\!-\!\!-\!\!-\!\!\bullet\ 2 \quad . \tag{165}$$

and the second term can be represented as the sum of two diagrams

$$\langle 0|\mathrm{T}\varphi_1\varphi_2\frac{(-i\lambda)}{4!}\int d^4y\varphi^4(y)|0\rangle = 3\left(\frac{-i\lambda}{4!}\right)\Delta_{12}\int d^4y\Delta_{yy}\Delta_{yy} \tag{166}$$

$$+\,12\left(\frac{-i\lambda}{4!}\right)\int d^4y\Delta_{1y}\Delta_{2y}\Delta_{yy} \tag{167}$$

$$= \quad 1 \bullet\!\!-\!\!-\!\!\bullet\ 2 \quad \overset{y}{\bigcirc} \quad +\,1 \bullet\!\!-\!\!-\!\!\underset{y}{\bigcirc}\!\!-\!\!\bullet\ 2 \quad .$$

The 3 and the 12 are the number of ways of pairing the field operators in Wick's theorem. For example the three different ways of pairing the φ_y's in the first term are

$$\overbrace{\varphi_y\varphi_y}\overbrace{\varphi_y\varphi_y} \tag{168}$$

$$\varphi_y\overbrace{\varphi_y\varphi_y\varphi_y}\varphi_y \tag{169}$$

$$\varphi_y\varphi_y\varphi_y\varphi_y\,. \tag{170}$$

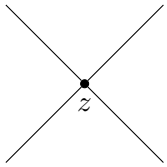
These first two terms in the expansion of the numerator are sufficient for us to guess how to go backwards from the diagram to the expression.

Position space Feynman rules for the numerator $\langle 0|\mathrm{T}\varphi_1\cdots\varphi_n e^{i\int d^4y\mathcal{L}_{int}}|0\rangle$: The numerator $\langle 0|\mathrm{T}\varphi_1\cdots\varphi_n e^{i\int d^4y\mathcal{L}_{int}}|0\rangle$ is the sum of all diagrams with n external points.

1. For each line $x \bullet\!\!-\!\!-\!\!\bullet\ y$ $= \Delta_F(x-y)$. Our Feynman diagrams were constructed by using this rule in the other direction.

2. For each external point $-\!\!-\!\!\bullet\ x$ $= 1$.

   We only include this rule to prepare you for particles of non-zero spin. For higher spin particles the rule for external points is non-trivial.

3. For each vertex 

$= -i\lambda \int d^4z$.

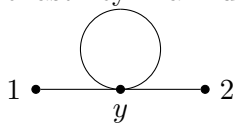
   Every factor of λ in the perturbative expansion of the exponential comes with an integral and four fields evaluated at the same point. After using Wick's theorem, these four fields become contracted. This contraction leads to four lines ending at the same point.

   More generally if we have a vertex where n_1 lines of type φ, n_2 lines of type ϕ, and n_3 lines of type ψ, then the vertex is associated with a factor of $-ig \int d^4z$ if $\mathcal{L}_{\text{int}}$ contains a term $\frac{-g}{n_1!n_2!n_3!}\varphi^{n_1}\phi^{n_2}\psi^{n_3}$. It is crucial (and standard) to normalize the Lagrangian in this way in order for the next Feynman rule to work.

4. Divide by the symmetry factor. We have accounted for everything except the overall constant. How to compute the symmetry factor is the subject of the next subsection.

7.1.1. Symmetry Factors

The last Feynman rule requires more explanation. The diagram



has coefficient

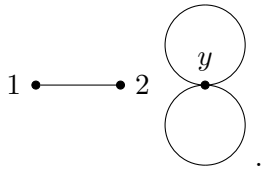
$$\frac{12}{4!} = \frac{1}{2} = \frac{1}{S}\,. \tag{171}$$

Here 12 is the number of Wick contractions, and the factor of $\frac{1}{4!}$ came from the interaction Lagrangian. S is the number of possible ways to exchange components of the diagram without changing the diagram, with external points held fixed[18]. In this case we can exchange the two ends of the internal line so $S = 2$. There are three types of symmetries to look for:

1. Exchanging ends of an internal line
2. Exchanging internal lines
3. Exchanging vertices or subdiagrams

We also get the correct numerical factor for the other $\mathcal{O}(\lambda)$ term in the numerator of the two point function

[18] In other words it is the order of the automorphism group.



This diagram has a symmetry factor

$$S = 2 \cdot 2 \cdot 2 = 8\,. \tag{172}$$

The first factor of 2 comes from exchanging the ends of the upper line, the second factor of 2 comes from exchanging the ends of the lower line, and the last factor of 2 comes from exchanging the lines. We get the same factor as before

$$\frac{1}{8} = \frac{3}{4!} \tag{173}$$

where 3 is the number of different Wick contractions and the denominator comes from the coefficient of the interaction Lagrangian $\frac{\lambda}{4!}$.

Computing symmetry factors is entirely equivalent to counting the number of Wick contractions. In practice most symmetry factors are 1 or 2, so computing the symmetry factor is often easier.

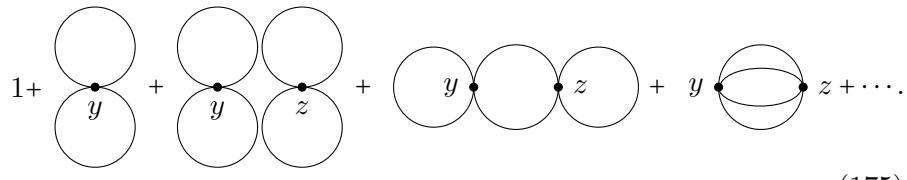
7.1.2. Vacuum Diagrams

Ultimately we are more interested in computing the Green's function than in the numerator or denominator of Dyson's formula. We will see that the rules for the numerator and the rules for the Green's function are similar. To determine the Feynman rules for the Green's function we first need to examine the denominator.

As an example let us look at the denominator of the two-point function

$$\langle 0|\mathrm{T}e^{i\int d^4x\mathcal{L}_{\mathrm{int}}}|0\rangle\,. \tag{174}$$

In φ^4 theory this is

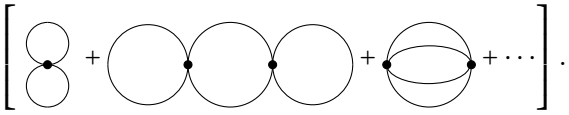
$\langle 0|\mathrm{T}e^{i\int d^4x\mathcal{L}_{int}}|0\rangle =$ 

(175)

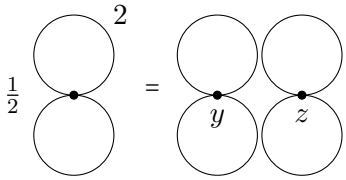
One of the order λ^2 terms contains the order λ term squared. It seems like we might be able to get both of these terms from an exponential. After all our original expression is an

exponential. A reasonable guess is

$$\text{denom} = \exp[\text{connected vacuum diagrams}] \tag{176}$$



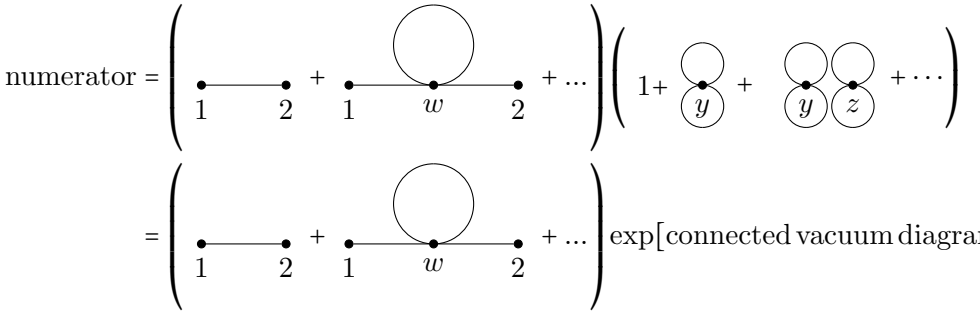
If we expand out the exponential we get a term that looks like the first order term squared



(177)

where the $\frac{1}{2}$ on the left hand side comes from expanding the exponential to second order. The symmetry factor on the right hand side has an extra factor of 2 from the $y \leftrightarrow z$ symmetry. Our guess passes this test. It can be shown that our guess is correct.

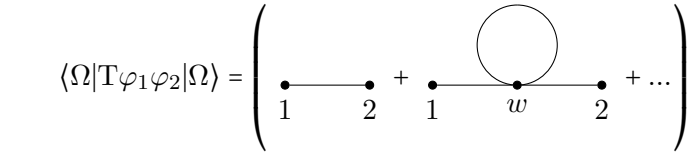
The terms in the numerator contain a part with external points times a part with vacuum diagrams



(178)

$$= \left(\quad \right) \exp[\text{connected vacuum diagrams}] .$$

The exponential of the connected vacuum diagrams cancels between the numerator and denominator so the two-point Green's function is



(179)

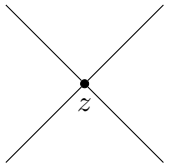
This cancellation of the vacuum diagrams between the numerator and denominator of Dyson's formula works for n-point functions as well.

The only difference between the Feynman rules for the numerator and the Feynman rules for the Green's function is that diagrams with vacuum subdiagrams do not contribute to

Green's functions. Since we are ultimately more interested in computing Green's functions than in computing the numerator, we reproduce the Feynman rules for position space Green's functions here for your reference.

Position space Feynman rules for $\langle\Omega|\mathrm{T}\varphi_1\dots\varphi_n|\Omega\rangle$:

The position-space n-point Green's function $\langle\Omega|\mathrm{T}\varphi_1\dots\varphi_n|\Omega\rangle$ is the sum of all diagrams with n external points and no vacuum subdiagrams.

1. For each line, $x \bullet\!\!-\!\!-\!\!\bullet\, y \;= \Delta_F(x-y)$
2. For each external point $-\!\!-\!\!\bullet\, x \;= 1$.
3. For each vertex  $= -i\lambda \int d^4z$.

   This rule needs to be modified depending on the theory. Some theories have multiple allowed vertices.
4. Divide by symmetry factor.

7.2. Momentum Space Feynman Rules

Feynman diagrams can also be used to compute matrix elements $i\mathcal{M}$. These matrix elements can then be used to compute cross sections or decay rates. In this subsection we will develop Feynman rules for matrix elements by using the LSZ reduction formula together with the Feynman rules for Green's functions that we developed in the previous subsection.

7.2.1. Connected Versus Disconnected Diagrams

As before we will look an example to deduce the Feynman rules. The simplest interesting scattering process is $2 \to 2$ scattering. To determine which Feynman diagrams contribute to the amplitude

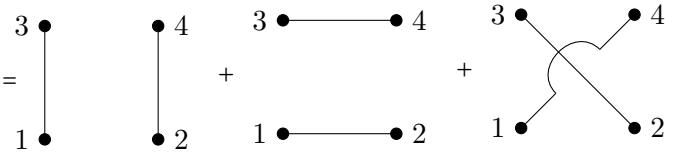
$$\langle \mathbf{p}_3, \mathbf{p}_4|S|\mathbf{p}_1, \mathbf{p}_2\rangle = \langle f|S|i\rangle \tag{180}$$

we can use the LSZ reduction formula

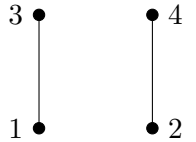
$$\begin{aligned}\langle f|S|i\rangle = i^4 \int d^4x_1 d^4x_2 d^4x_3 d^4x_4 e^{i(p_1\cdot x_1 + p_2\cdot x_2 - p_3\cdot x_3 - p_4\cdot x_4)} \\ \cdot (\partial_1^2 + m^2)(\partial_2^2 + m^2)(\partial_3^2 + m^2)(\partial_4^2 + m^2)\langle\Omega|\mathrm{T}\varphi_1\varphi_2\varphi_3\varphi_4|\Omega\rangle\end{aligned} \tag{181}$$

together with the Feynman diagrams that contribute to the four-point Green's function.

At leading order the four-point Green's function is

$\langle\Omega|\mathrm{T}\varphi_1\varphi_2\varphi_3\varphi_4|\Omega\rangle =$  $+\mathcal{O}(\lambda).$ (182)

where time flows upward. Using the position space Feynman rules the first term is

 $= \Delta_F(x_1 - x_3)\Delta_F(x_2 - x_4)$ (183)

Let's isolate some parts of the LSZ reduction formula. Define

$$F(x_1 - x_3) \equiv (\partial_1^2 + m^2)(\partial_3^2 + m^2)\Delta_F(x_1 - x_3)\,. \tag{184}$$

We can perform the x_1 and x_3 integrals in the LSZ formula using new variables

$$\begin{aligned} x_{13} &= x_1 + x_3, \;\; p_{13} = \frac{p_1 + p_3}{2} \\ \bar{x}_{13} &= x_1 - x_3, \;\; \bar{p}_{13} = \frac{p_1 - p_3}{2}\,. \end{aligned} \tag{185}$$

The x_1 and x_3 integrals become

$$\begin{aligned} \int d^4x_1 d^4x_3 e^{i(p_1\cdot x_1 - p_3\cdot x_3)} F(x_1 - x_3) &= \frac{1}{2^4}\int d^4x_{13} d^4\bar{x}_{13} e^{i(\bar{p}_{13}\cdot x_{13} + p_{13}\cdot\bar{x}_{13})} F(\bar{x}_{13}) \\ &= \frac{1}{2^4}\int d^4x_{13} e^{i\bar{p}_{13}x_{13}}\tilde{F}(p_{13}) \\ &= (2\pi)^4\delta^{(4)}(p_1 - p_3)\tilde{F}\left(\frac{p_1 + p_3}{2}\right). \end{aligned} \tag{186}$$

The factor of $\frac{1}{2^4}$ comes from the Jacobian of the change of variables. The x_2 and x_4 integrals work similarly. Thus the scattering amplitude for this term is proportional to

$$\langle f|S|i\rangle_{||} = ||\text{term in } \langle f|S|i\rangle \propto \delta^{(4)}(p_1 - p_3)\delta^{(4)}(p_2 - p_4) \tag{187}$$

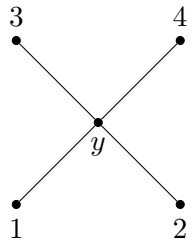
There is no scattering! This is the same for the cross diagram. The = diagram vanishes, as

$$\langle f|S|i\rangle_{=} \propto \delta^{(4)}(p_1 + p_2)\delta^{(4)}(p_3 + p_4) \tag{188}$$

but the energy of particles 1 and 2 is positive $p_1^0 + p_2^0 \geq 2m > 0$. The same argument works for other diagrams that are not fully connected.

Only fully connected diagrams contribute to two-to-two scattering. If more particles are scattering there can be disconnected diagrams that do contribute, but they are uninteresting as they do not involve all of the particles interacting. Disconnected diagrams have extra deltas compared to fully connected diagrams. The S-matrix factorizes into a product of sums of connected diagrams. Connected and disconnected diagrams never interfere due to the extra deltas in the disconnected diagrams.

The lowest order fully connected diagram is



$$= -i\lambda \int d^4y \Delta_{1y}\Delta_{2y}\Delta_{3y}\Delta_{4y}\,. \tag{189}$$

The Klein-Gordon operators in the LSZ reduction formula (181) turn these propagators into deltas because the Feynman propagator is a Green's function of the Klein-Gordon operator

$$(\partial_1^2 + m^2)\Delta_{1y} = -i\delta^{(4)}(x_1 - y)\,. \tag{190}$$

We can use these deltas to do all four x integrals in the LSZ reduction formula (181) leaving us with one delta

$$\begin{aligned}\langle f|S|i\rangle_\times &= -i\lambda \int d^4y e^{i(p_1+p_2-p_3-p_4)y} \\ &= -i\lambda(2\pi)^4\delta^{(4)}(p_1+p_2-p_3-p_4)\end{aligned} \tag{191}$$

The momentum conserving delta $(2\pi)^4\delta^{(4)}(p_1+p_2-p_3-p_4)$ occurs in all fully connected $2 \to 2$ diagrams. Our example calculation has reproduced the 4-momentum conserving delta that we expected

$$\langle f|S|i\rangle = i\mathcal{M}(i\to f)(2\pi)^4\delta^{(4)}\left(\sum_{\text{incoming}} p - \sum_{\text{outgoing}} p\right) \tag{192}$$

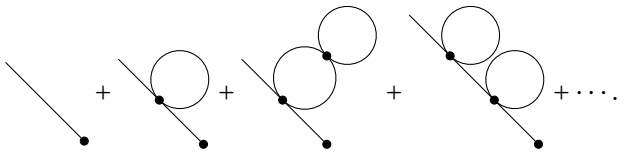
where $i\mathcal{M}(i \to f)$ is the matrix element. The contribution to the matrix element from the cross diagram is

$$i\mathcal{M}(i\to f)_\times = -i\lambda \tag{193}$$

The Klein-Gordon operators $(\partial^2 + m^2)$ eliminated the external propagators. These Klein-Gordon operators in the LSZ reduction formula require us to treat external lines in Feynman diagrams differently from internal lines when we use the Feynman rules for matrix elements.

7.2.2. Amputation

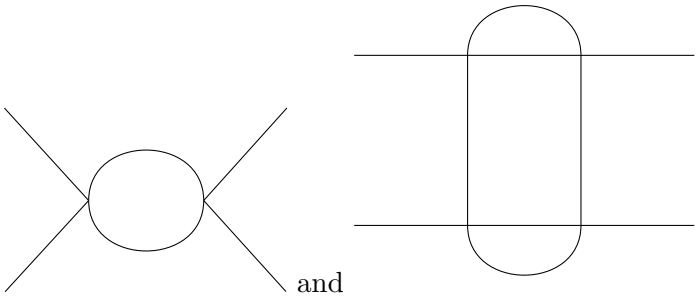
The m^2 in the LSZ reduction formula represents the physical mass. Corrections to the external legs represent evolution of the momentum eigenstates in the free theory to the momentum eigenstates of the interacting theory $|\mathbf{p}\rangle_0 \to |\mathbf{p}\rangle$. The operator $(\partial^2 + m^2)$ kills the propagators in the external legs with corrections. In other words a one-particle state in the interacting theory can be represented by a sum of Feynman diagrams

one particle state =  (194)

Colloquially speaking, a particle is never alone in quantum field theory. Instead it is constantly interacting with virtual particles.

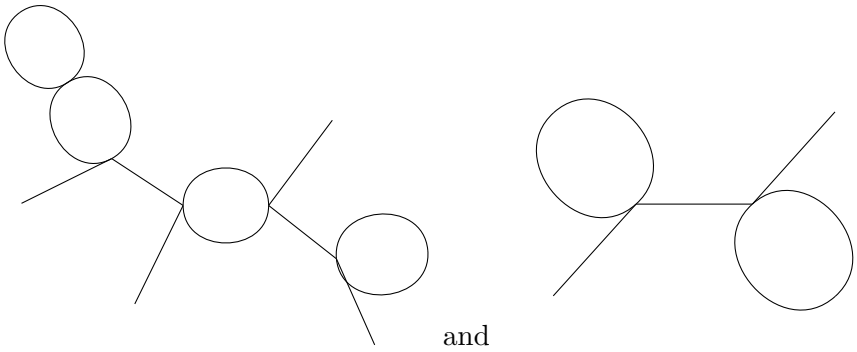
These diagrams represent corrections that can change the location of the pole in the propagator as we will see in more detail in the next chapter. After we account for the change in the location of the pole the LSZ formula will eliminate these diagrams.

For example the diagrams



and

are amputated, and the diagrams



and

are not amputated.

7.2.3. Momentum Conservation

Additional deltas arise in higher-order fully connected diagrams. These deltas come from performing the position integrals that accompany each factor of the coupling constant as in (191). These deltas impose momentum conservation at each vertex. For fully connected diagrams, there are sufficiently many momentum integrals to eliminate the momentum-conserving deltas at each vertex. In fact, some momentum integrals will remain in the contribution to the matrix element $i\mathcal{M}$.

To see how many momentum integrals will remain in the matrix element, consider a fully connected $2 \to 2$ diagram in φ^4 theory with V vertices and E edges or lines.

Each vertex is accompanied by a spacetime position integral $\int d^4x$. Each internal line is associated with a Feynman propagator

$$\Delta_F(x-y) = \int d^4k \frac{ie^{ik\cdot(x-y)}}{k^2 - m^2 + i\epsilon}\,. \tag{195}$$

The complex exponentials from the Feynman propagator combine with the spacetime position integrals to yield momentum conserving deltas. The external lines (of which there are four for $2 \to 2$ scattering) are amputated so do not contribute any momentum integrals. The internal lines each contribute a momentum integral. One of the deltas[19] conserves overall momentum. This overall momentum conserving delta is part of the scattering amplitude $\langle f|S|i\rangle$ but not the matrix element $i\mathcal{M}$.

In total the diagram's contribution to the matrix element has $V-1$ momentum conserving deltas and $E-4$ momentum integrals. The number of remaining momentum integrals is

$$\text{number of momentum integrals} = E - 4 - (V-1) = E - V - 3\,. \tag{196}$$

Euler's formula states that the Euler characteristic $\chi = V - E + L$ is a constant for Feynman diagrams with a fixed number of external lines where L is the number of independent loops. For the Feynman diagrams of interest (connected diagrams with four external lines) $\chi = -3$ so

$$L = E - V - 3\,. \tag{197}$$

Combing (196) and (197) tells us that the number of remaining momentum integrals is equal to the number of independent loops.

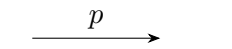
7.2.4. Momentum Space Feynman Rules for Scattering Amplitudes

We now collect the lessons of the previous subsections and write down the Feynman rules for the matrix element in momentum space. We saw that the matrix element[20] $i\mathcal{M}$ is the sum of all completely connected, amputated diagrams. For each diagram, we can write down the associated expression using the following rules:

[19] Or combination of deltas.

[20] More precisely, the interesting part of the matrix element (that involves all particles interacting).

1. For each internal line



$$= \frac{i}{p^2 - m^2 + i\epsilon}\,. \tag{198}$$

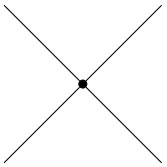
   The best practice is to label the direction of momentum on each line with an arrow next to the line. We reserve arrows on the lines to indicate the flow of charge.

2. For each external line

$$= 1\,. \tag{199}$$

   (This rule will change for fermions and gauge bosons.)

3. For each vertex



$$= -i\lambda\,. \tag{200}$$

   (This rule is specific to $\lambda\varphi^4$ theory.)

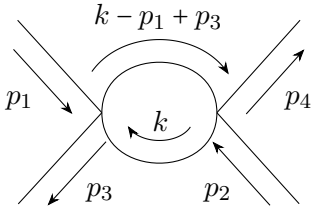
4. Impose momentum conservation at each vertex. No momentum conserving deltas should be present in the matrix element.

5. Integrate over any undetermined momenta. There should be one undetermined momentum per loop.

6. Divide by the symmetry factor.

As we saw in the previous subsections, these rules follow from the LSZ reduction formula and the Feynman rules for the position space Green's functions.

For example



$$= \frac{(-i\lambda)^2}{2} \int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2 - m^2 + i\epsilon} \frac{i}{(k + p_2 - p_4)^2 - m^2 + i\epsilon}\,. \tag{201}$$

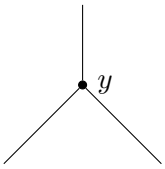
8. Renormalization

To consistently make predictions using perturbation theory beyond leading order, we must renormalize our theory.

Suppose we are interested in a theory in which there is a cubic self interaction for a real scalar field. The naive Lagrangian describing this theory is

$$\mathcal{L}_{\text{naive}} = \frac{1}{2}(\partial\varphi)^2 - \frac{1}{2}m^2\varphi^2 + \frac{1}{3!}g\varphi^3 \,. \tag{202}$$

We expect this theory to have two physical parameters, the mass and the strength of the coupling of the particles. In the previous chapters we have learned how to make predictions at leading order in the coupling. To compute the leading term in the perturbative expansion of the Green's functions we can use the Feynman rule for the vertex in position space



$$= ig \int d^4y \,. \tag{203}$$

We cannot consistently use this form of the Lagrangian beyond leading order in the LSZ reduction formula as presented in chapter 5. It is inconsistent with the assumption we made that the field and ladder operators are related the same way in the free and interacting theory. Our assumptions in deriving the LSZ reduction formula imply that

$$\begin{aligned} \langle\Omega|\varphi(x)|\Omega\rangle &= 0 \\ \langle\mathbf{k}|\varphi(x)|\Omega\rangle &= e^{ik\cdot x} \,. \end{aligned} \tag{204}$$

As we saw in chapter 5, to impose these conditions we have to shift and rescale (and rename). Our Lagrangian becomes

$$\mathcal{L} = \frac{1}{2}Z_\varphi(\partial\varphi)^2 - \frac{1}{2}Z_m m^2\varphi^2 + \frac{1}{3!}Z_g g\varphi^3 + Y\varphi \,. \tag{205}$$

The Lagrangian can be separated into a free part and an interaction part

$$\mathcal{L} = \mathcal{L}_{\text{free}} + \mathcal{L}_{\text{int}} + \mathcal{L}_{\text{ct}} \tag{206}$$

where

$$\mathcal{L}_{\text{free}} = \frac{1}{2}(\partial\varphi)^2 - \frac{1}{2}m^2\varphi^2 \tag{207}$$

and

$$\mathcal{L}_{\text{int}} = \frac{1}{3!}Z_g g\varphi^3 + \mathcal{L}_{\text{ct}} \,. \tag{208}$$

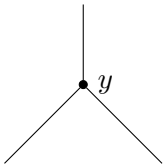
We have introduced the counterterm Lagrangian

$$\mathcal{L}_{\text{ct}} = \frac{1}{2}(Z_\varphi - 1)(\partial\varphi)^2 - \frac{1}{2}(Z_m - 1)m^2\varphi^2 + Y\varphi \,. \tag{209}$$

Counterterms are the terms we are forced to add.[21] We should recover the free theory in the limit $g \to 0$ so at leading order Y vanishes and the Z_i are all unity

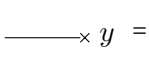
$$\begin{aligned} Y &= 0 + \mathcal{O}(g) \\ Z_i &= 1 + \mathcal{O}(g^2)\,. \end{aligned} \tag{210}$$

The first order corrections to the Z_i vanish as we will see when we calculate them. Compared to our naive theory, the Feynman rule for the vertex changes



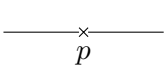
$$= iZ_g g \int d^4y\,. \tag{211}$$

We have a new vertex since we have a linear term $Y\varphi$ in the interaction Lagrangian. In position space, the Feynman rule for this new vertex is



$$= iY \int d^4y\,. \tag{212}$$

The line ends at the point y, but y is an internal point. There is also a new vertex where two lines meet that is easier in momentum space

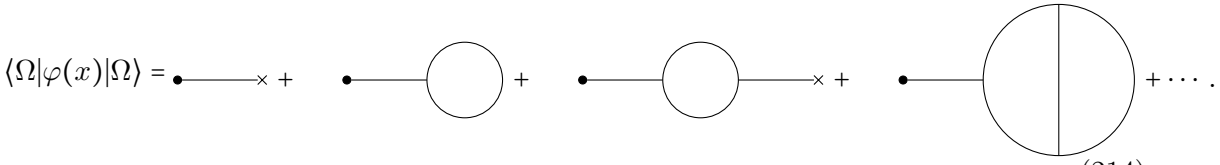


$$= ip^2(Z_\varphi - 1) - i(Z_m - 1)m^2\,. \tag{213}$$

In the following subsections we will determine the Z_i and Y at next-to-leading order. We need four conditions to determine these four parameters. Two of these conditions will come from the assumptions we made in the LSZ reduction formula. The other two conditions will be conditions that relate the two physical parameters (mass and coupling) of the theory to the parameters that appear in the Lagrangian.

8.1. One-Point Function

The value of Y can be determined by examining the one-point function



(214)

[21] Sometimes $Z_g g/Z_\varphi^{3/2}$ is called the bare coupling and $Z_m m^2/Z_\varphi$ is called the bare mass squared. The renormalized coupling is g and the renormalized mass squared is m^2.

These diagrams are known as tadpole diagrams. In deriving the LSZ formula we assumed that the one-point function vanishes. The diagrams that contribute to the one-point function must sum to zero at each order. At $\mathcal{O}(g)$ the first two diagrams contribute

$$\begin{aligned}\langle\Omega|\varphi(x)|\Omega\rangle &= iY\int d^4y\Delta_F(x-y)+\frac{1}{2}(ig)\int d^4y\Delta_F(x-y)\Delta_F(y-y)+\mathcal{O}(g^2) \\ &= \left(iY+\frac{1}{2}ig\Delta_F(0)\right)\int d^4y\Delta_F(x-y)+\mathcal{O}(g^2)\end{aligned} \tag{215}$$

where we did not include the factor of Z_g for the vertex because it is unity up to second order [22]. For the one-point function vanish at order g, Y must be

$$Y = -\frac{1}{2}g\Delta_F(0)+\mathcal{O}(g^3)\,. \tag{216}$$

There is a small problem. The Feynman propagator evaluated at zero diverges

$$\Delta_F(0) = \int\frac{d^4p}{(2\pi)^4}\frac{i}{p^2-m^2+i\epsilon}e^{-ip\cdot 0}\,. \tag{217}$$

We can proceed by introducing a regulator, or a prescription for making the integral finite like we did when we computed the Casimir force. The precise method we use to regulate the theory will not affect our final result. In this course we will mostly regulate our theories by imposing a UV cutoff Λ.[23] In the next subsection we evaluate the integral with Pauli-Villars regularization in the limit $\Lambda^2 \gg m^2$. The result is

$$\Delta_F(0) = \frac{\Lambda^2}{16\pi^2}\,. \tag{219}$$

This result diverges in the limit $\Lambda\to\infty$, but it cancels out of all physically measurable quantities. If we impose

$$\langle\Omega|\varphi(x)|\Omega\rangle = 0 \tag{220}$$

then the sum of all diagrams with tadpole diagrams will cancel at each order in perturbation theory. Therefore there is no practical reason to compute any diagram with a tadpole subdiagram. For pedagogical purposes we compute the tadpole diagram in the next section. The integrals we care about[24] can be computed using similar techniques.

[22] We only used that Z_g is unity at zeroth order in g.

[23] Other regulators make some calculations easier (or possible at all - as we will see shortly sometimes regulators fail). We could modify the propagator to improve its UV behavior

$$\Delta_F(0)\to\int\frac{d^4p}{(2\pi)^4}\frac{i}{p^2-m^2+i\epsilon}\left(\frac{\Lambda^2}{p^2-\Lambda^2+i\epsilon}\right)^2\,. \tag{218}$$

This regularization procedure is known as Pauli-Villars regularization. Dimensional regularization involves changing the dimension of spacetime (to a complex number!) so that the integral converges.

[24] Similar integrals will arise when we compute the Z's.

8.1.1. Wick Rotation

The divergent integral (217) would be easier to evaluate in hyperspherical coordinates. Unfortunately the minus signs in the metric prevent us from directly changing to hyperspherical coordinates. To eliminate these pesky minus signs, we will choose a different contour to evaluate the p^0 integral. Instead of integrating along the real p^0 axis, we will integrate along the imaginary p^0 axis. Without passing through any poles, we can rotate the contour counterclockwise by $\frac{\pi}{2}$ so that p^0 is integrated from $-i\infty$ to $+i\infty$. This rotation is known as a Wick rotation. The $i\epsilon$ determines which side of the real axis the poles are on and hence which way we should rotate the contour. The integral along the imaginary axis is equal to the integral along the real axis provided the integrand goes to zero sufficiently quickly for large $|p^0|$. After Wick rotation, we can define a Euclidean 4-momentum p_E

$$p_E^0 = -ip^0; \qquad \mathbf{p}_E = \mathbf{p} \tag{221}$$

so $p_E^2 = -p^2$.

A Wick rotation is not justified for the original integral (217) because the integrand does not go to zero sufficiently quickly as $|p^0| \to \infty$. Instead we can consider the regulated integral

$$\Delta_F(0;\Lambda) = \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2 + i\epsilon} \left(\frac{\Lambda^2}{p^2 - \Lambda^2 + i\epsilon} \right)^2 . \tag{222}$$

In the limit $\Lambda \to \infty$ the extra factor we inserted becomes unity. The extra factor improves the convergence of the integrand for large $|p^0|$ and allows us to Wick rotate. After Wick rotation the regulated integral becomes

$$\Delta_F(0;\Lambda) = \int \frac{d^4p_E}{(2\pi)^4} \frac{1}{p_E^2 + m^2} \left(\frac{\Lambda^2}{p_E^2 + \Lambda^2} \right)^2 . \tag{223}$$

In hyperspherical coordinates this integral becomes

$$\begin{aligned} \Delta_F(0;\Lambda) &= \int d\Omega_4 \int_0^\infty \frac{|p_E|^3 d|p_E|}{(2\pi)^4} \frac{1}{|p_E|^2 + m^2} \left(\frac{\Lambda^2}{|p_E|^2 + \Lambda^2} \right)^2 \\ &= 2\pi^2 \frac{\Lambda^4}{(2\pi)^4} \frac{1 + \frac{m^2}{\Lambda^2 - m^2} \log\left(\frac{m^2}{\Lambda^2}\right)}{2(\Lambda^2 - m^2)} \\ &\approx \frac{\Lambda^2}{16\pi^2} \end{aligned} \tag{224}$$

where $\int d\Omega_4 = 2\pi^2$ is the surface area of a 4-sphere[25]. The final line is valid in the limit $\Lambda^2 \gg m^2$.

[25] The radial integral can be done in Mathematica or by using a table of integrals if you are several hundred years old.

8.2. Propagator

Our second condition $\langle \mathbf{k}|\varphi(x)|\Omega\rangle = e^{ik\cdot x}$ is related to the propagator since by assumption $\langle \mathbf{k}|$ and $\lim_{t\to\pm\infty}\varphi(x)|\Omega\rangle$ are one particle states. In this section we will discuss two important general results for the propagator. These results will allow us to determine Z_m and Z_φ at leading order.

For the most part it will be more convenient to work in position space

$$G(x,x') = \langle\Omega|T\varphi(x)\varphi(x')|\Omega\rangle \tag{225}$$

or in momentum space

$$\begin{aligned} G(p,p') &= \int d^4x \int d^4x' G(x,x') e^{i(p\cdot x + p'\cdot x')} \\ &= (2\pi)^4 \delta^{(4)}(p+p') G_2(p) \end{aligned} \tag{226}$$

instead of the hybrid position-momentum space equation $\langle k|\varphi(x)|\Omega\rangle = e^{ik\cdot x}$.

8.2.1. Källén-Lehmann Spectral Representation

The Källén-Lehmann spectral representation is an exact expression for the propagator of an interacting theory. We will sketch the ideas behind the derivation of the Källén-Lehmann equation.

We will assume that:

- There are no bound states.
- $\langle\Omega|\varphi(x)|\Omega\rangle = 0$
- $\langle \mathbf{k}|\varphi(x)|\Omega\rangle = e^{ik\cdot x}$

The last two assumptions are a consequence of the assumptions we made when we derived the LSZ reduction formula.[26]

The identity can be written as a sum and integral over the zero-, one-, and multi-particle states

$$\mathbf{1} = |\Omega\rangle\langle\Omega| + \int \frac{d^3k}{(2\pi)^3 2E_{\mathbf{k}}} |\mathbf{k}\rangle\langle\mathbf{k}| + \sum\!\!\!\!\!\!\int_\sigma \int \frac{d^3k}{(2\pi)^3 2E_{\mathbf{k}}} |\mathbf{k},\sigma\rangle\langle\mathbf{k},\sigma| \tag{227}$$

where $\mathbf{k}$ is the total 3-momentum and σ denotes the other discrete and continuous parameters necessary to label the states. If we insert the above resolution of the identity in between $\varphi(x)$ and $\varphi(y)$ in (with $x^0 > y^0$)

$$\begin{aligned} \langle\Omega|\varphi(x)\varphi(y)|\Omega\rangle = \langle\Omega|\varphi(x)|\Omega\rangle\langle\Omega|\varphi(y)|\Omega\rangle &+ \int \frac{d^3k}{(2\pi)^3 2E_{\mathbf{k}}} \langle\Omega|\varphi(x)|\mathbf{k}\rangle\langle\mathbf{k}|\varphi(y)|\Omega\rangle \\ &+ \sum\!\!\!\!\!\!\int_\sigma \int \frac{d^3k}{(2\pi)^3 2E_{\mathbf{k}}} \langle\Omega|\varphi(x)|\mathbf{k},\sigma\rangle\langle\mathbf{k},\sigma|\varphi(y)|\Omega\rangle \end{aligned} \tag{228}$$

[26] These assumptions can be relaxed at the expense of complicating our final result.

then the zero-particle term vanishes since by assumption $\langle\Omega|\varphi(x)|\Omega\rangle = 0$. The momentum integrals in the one- and multi-particle state terms can be rewritten using the identity from section 3.4

$$\int \frac{d^3k}{(2\pi)^3 2E_{\mathbf{k}}} = \int \frac{d^4k}{(2\pi)^4}(2\pi)\delta(k^2 - M^2)\Theta(k^0) \tag{229}$$

where M is the (physically observable!) invariant mass of the one- or multi-particle state. Since $|\mathbf{k},\sigma\rangle$ are momentum eigenstates and the vacuum is translation invariant

$$\langle\mathbf{k},\sigma|\varphi(x)|\Omega\rangle = e^{ik\cdot x}\langle\mathbf{k},\sigma|\varphi(0)|\Omega\rangle\,. \tag{230}$$

Combining these observations we see that the two point function is proportional to the spectral density $\rho(s)$ (with $s = k^2$)

$$\langle\Omega|\varphi(x)\varphi(y)|\Omega\rangle \propto \rho(s) = 2\pi\delta(s - m_{\text{ph}}^2) + \sum\!\!\!\!\!\!\int_{\sigma} 2\pi\delta(s - M_\sigma^2)|\langle\mathbf{k},\sigma|\varphi(0)|\Omega\rangle|^2 \tag{231}$$

where m_{ph} is the physical mass of a single particle and $M_\sigma > 2m_{\text{ph}}$ is the invariant mass of the multi-particle state. After similar manipulations[27] to those we did when we calculated the Feynman propagator in the free theory it can be shown that the exact propagator in momentum space is given by

$$G_2(k) = \frac{i}{k^2 - m_{\text{ph}}^2 + i\epsilon} + \int_{4m_{\text{ph}}^2}^{\infty} \frac{ds}{2\pi}\rho(s)\frac{i}{k^2 - s + i\epsilon}\,. \tag{232}$$

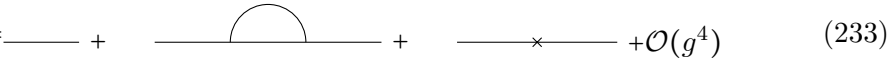
The first term is the free propagator[28]. The second term contains the contributions of the multiparticle states.

The Källén-Lehmann spectral representation guarantees that exact propagator has a pole at the physical mass. The residue at the pole at $k^2 = m_{\text{ph}}^2$ is i. These properties of the exact propagator must hold at each order in perturbation theory.

The Klein-Gordon operators in the LSZ reduction formula also involve the physical mass. Together the LSZ and Källén-Lehmann formulae justify the amputation procedure described in the previous chapter.

8.2.2. Propagator at Order g^2

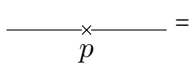
The momentum space propagator at order g^2 is

$G_2(p) =$  $+\mathcal{O}(g^4)$ (233)

[27] We need to combine the expressions for $x^0 > y^0$ with those for $y^0 > x^0$ to get the time-ordered two point function. Then we can write the result as a contour integral, and finally take the Fourier transform.

[28] In the free theory there is no difference between the mass parameter in the Lagrangian and the physical mass. These two masses are not necessarily the same in an interacting theory.

Note that the third diagram uses the new counterterm vertex



$$= i(Z_\varphi - 1)p^2 - i(Z_m - 1)m^2 \,. \tag{234}$$

We will work in d spacetime dimensions for now. As we will see as we compute the Feynman diagrams, making a choice other than $d = 4$ will simplify some calculations.

Using the momentum space Feynman rules[29] the propagator is

$$\begin{aligned} G_2(p) =& \frac{i}{p^2 - m^2 + i\epsilon} + \frac{(-ig)^2}{2}\left(\frac{i}{p^2 - m^2 + i\epsilon}\right)^2 \int \frac{d^dk}{(2\pi)^d} \frac{i}{k^2 - m^2 + i\epsilon} \frac{i}{(p-k)^2 - m^2 + i\epsilon} \\ &+ \left(i(Z_\varphi - 1)p^2 - i(Z_m - 1)m^2\right)\left(\frac{i}{p^2 - m^2 + i\epsilon}\right)^2 . \end{aligned} \tag{235}$$

In the second term Z_g does not appear because the corrections to Z_g in the second term contribute to the propagator at order g^4. The Källén-Lehmann spectral representation guarantees that the propagator has a pole at the physical mass and determines the residue at this pole. We can use these conditions to determine Z_φ and Z_m.

To determine Z_φ and Z_m we need to compute the integral in the second term

$$I(p^2) = \int \frac{d^dk}{(2\pi)^d} \frac{i}{k^2 - m^2 + i\epsilon} \frac{i}{(p-k)^2 - m^2 + i\epsilon} \,. \tag{236}$$

The second propagator involves $(p-k)^2$ instead of the square of the vector integration variable so does not simplify as nicely in spherical coordinates. To fix this issue we can introduce a Feynman parameter x by using the identity[30]

$$\frac{1}{AB} = \int_0^1 dx \frac{1}{(xA + (1-x)B)^2} \,. \tag{237}$$

Choosing $A = (p-k)^2 - m^2 + i\epsilon$ and $B = k^2 - m^2 + i\epsilon$ the denominator becomes

$$D = (x(p-k)^2 + (1-x)k^2 - m^2 + i\epsilon)^2 = (k^2 - 2xp\cdot k + xp^2 - m^2 + i\epsilon)^2 \,. \tag{238}$$

If we complete the square the integral will be easier to compute in spherical coordinates

$$D = \left((k - xp)^2 + x(1-x)p^2 - m^2 + i\epsilon\right) \,. \tag{239}$$

[29] Note that even though we are working in momentum space, we are computing a Green's function so we do not need to amputate.

[30] This identity can be generalized to combine more than two denominators.

Making the change of variables[31] $q = k - xp$ and Wick rotating[32] $q \to q_E$ the integral becomes

$$I(p^2) = -i \int_0^1 dx \int \frac{d^d q_E}{(2\pi)^d} \frac{1}{(q_E^2 + x(1-x)p^2 + m^2)^2} \tag{240}$$

For $d = 2$ we can evaluate this integral to get

$$I(p^2) = -i \int_0^1 dx \frac{1}{4\pi} \frac{1}{x(1-x)p^2 + m^2} \,. \tag{241}$$

This integral is finite and we can use it to determine Z_φ and Z_m by demanding that the pole and residue of $G_2(p)$ is in the right place. Explicitly we need to require

$$-\frac{(ig)^2}{2} I(m^2) + i(Z_\varphi - 1)m^2 - i(Z_m - 1)m^2 = 0 \tag{242}$$

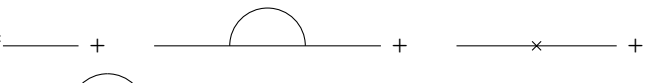
and

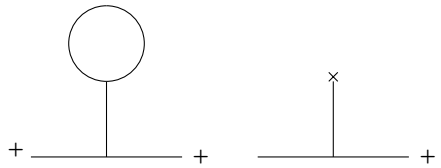
$$-\frac{(ig)^2}{2} I'(m^2) + i(Z_\varphi - 1) = 0 \tag{243}$$

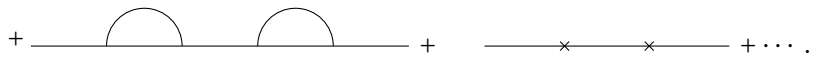
to ensure the pole and residue are correct.

8.2.3. Propagator at All Orders

If we continue our calculation of the propagator to higher order, the momentum space propagator in terms of Feynman diagrams is

$G_2(p) =$  (244)

 (245)

 (246)

The tadpole diagrams on the second line cancel. The diagrams on the third line can be constructed by connecting together copies of the diagrams on the first line. If we can compute the diagrams on the first line, we have done most of the work in computing the

[31] Changing variables in linearly divergent integrals can change the value of the integral. This fact is important in the computation of anomalies (quantum effects that spoil classical symmetries) using Feynman diagrams.

[32] This Wick rotation is justified when $d < 4$.

diagrams on the third line. Extending this idea, the propagator is a sum of one-particle irreducible diagrams connected by propagators

$G_2(p) =$ 

(247)

One particle irreducible diagrams are the ones that are still connected if any line is cut.

If we denote the contribution from all one-particle irreducible diagrams as $-i\Sigma(p^2)$ (Σ is known as the self-energy[33]) the propagator becomes a geometric series

$$\begin{aligned} G_2(p) &= \frac{i}{p^2-m^2+i\epsilon} + \left(\frac{i}{p^2-m^2+i\epsilon}\right)^2 (-i\Sigma(p^2)) + \left(\frac{i}{p^2-m^2+i\epsilon}\right)^3 (-i\Sigma(p^2))^2 + \dots \\ &= \frac{i}{p^2-m^2+i\epsilon} \sum_{n=0}^{\infty} \left(\frac{\Sigma(p^2)}{p^2-m^2+i\epsilon}\right)^n \\ &= \frac{i}{p^2-m^2+i\epsilon} \frac{1}{1-\frac{\Sigma(p^2)}{p^2-m^2+i\epsilon}} \\ &= \frac{i}{p^2-m^2-\Sigma(p^2)+i\epsilon}\,. \end{aligned} \tag{248}$$

This result is valid to all orders in perturbation theory in any theory!

Our two results for the propagator should agree with the Källen-Lehmann formula which tells us the propagator has an isolated pole at $p^2 = m_{\text{ph}}^2$ with residue i.

If we want[34] $m^2 = m_{\text{ph}}^2$ then $\Sigma(p^2)$ and its derivative must vanish at $p^2 = m^2$:

$$\Sigma(m^2) = 0 = \Sigma'(m^2) \tag{249}$$

We can use these conditions to fix Z_φ and Z_m at higher order in g by a generalization of the procedure we used in the previous subsection.

8.3. Vertex

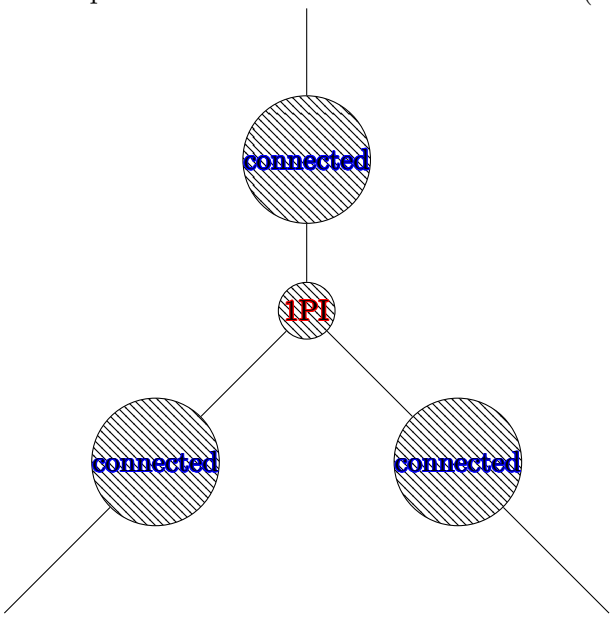
So far we have 3 conditions on 4 parameters (two from the Källen-Lehmann formula and one from tadpole cancellation). We need one more condition. We could demand some cross section or decay rate takes a particular value with a particular choice of momentum. Alternatively we could impose a condition on a Green's function with some particular choice of momentum. Both of these methods are more work than necessary.

The ideas of the last section generalize to Green's functions with more than two points. Every Feynman diagram is built from propagators and irreducible diagrams. We already

[33] The self-energy is closely related to the irreducible two-point function discussed later.

[34] This choice makes the physical interpretation clearest. Other choices make some computations easier.

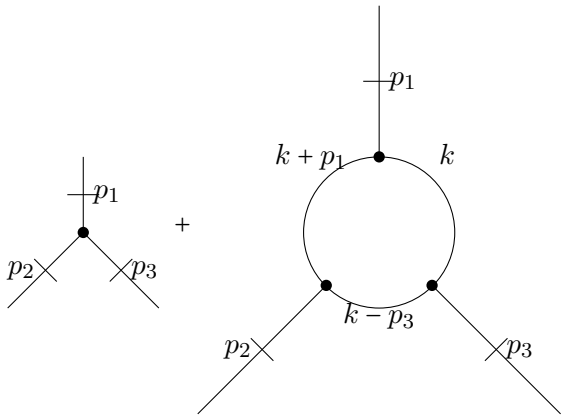
know how to compute the propagator. We need to compute the irreducible diagrams so that we understand all of the building blocks that we can use to construct Feynman diagrams. The 3-point Green's function is of the form (after cancellation of the tadpole diagrams)



with a three-point one-particle irreducible part in the middle, and connected 2-point functions on each of the legs. We already studied the connected 2-point function (or exact propagator) in the previous section.

An easier method is to impose a condition on the irreducible 3-point function. The irreducible three-point function at next-to-leading order is

$$\hat{\Gamma}_3(p_1,p_2,p_3) = \quad + \qquad (250)$$



The momenta on the external lines flow inward. The momenta in the loop flow counter clockwise.

Cutting a propagator with a bar represents applying the inverse propagator so

$$= 1 \tag{251}$$

Using the Feynman rules, the irreducible three-point function in momentum space at order g^3 is

$$\hat{\Gamma}_3(p_1,p_2,p_3) = iZ_g g + (ig)^3 \int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2-m^2+i\epsilon} \frac{i}{(k-p_1)^2-m^2+i\epsilon} \frac{i}{(k+p_3)^2-m^2+i\epsilon} \tag{252}$$

where we used $Z_g = 1+\mathcal{O}(g^2)$ in the second term. Note that the integral is finite. It can be evaluated using the techniques we have discussed (Feynman parameters and Wick rotation), although the computation is more involved than the ones we have already computed. In QFT I and II we will compute several irreducible n-point functions and use them to renormalize several scalar quantum field theories.

The irreducible three-point function is one of the building blocks for the Feynman diagrams. For example it will show up in the diagrams for two-to-two scattering in φ^3 theory. In a general context, the momenta in the three-point irreducible function will be off-shell (that is they will not satisfy $p^2 = m^2$). We can choose to define the coupling g by imposing that the irreducible three-point function is equal to the coupling with some choice of reference momenta

$$\hat{\Gamma}_3(p_1^{\text{ref}},p_2^{\text{ref}},p_3^{\text{ref}}) = g \tag{253}$$

where

$$(p_1^{\text{ref}})^2 = (p_3^{\text{ref}})^2 = \mu^2 \tag{254}$$

for some constant μ. We do not have to choose $\mu^2 = m^2$.

Our choice of μ was arbitrary. This fact has deep implications which we will explore further in the QFT I and II.

9. Future Directions

In this course we have developed scalar quantum field theory in the canonical formulation. We have focused on how to use quantum field theory to compute S-matrix elements. The techniques we have developed have many other applications.

Quantum field theory is a vast subject and some of the many interesting topics and applications that will be explored in future courses include[35]:

- Quantization of spinor and vector fields in QFT I

[35] This list is not exhaustive.

- Quantization of theories with abelian (QFT I) or non-abelian gauge symmetries[36] (QFT II)
- Functional integral quantization[37] in QFT II
- Conformal field theory[38] in Quantum Fields and Strings
- Anomalies[39] in Quantum Fields and Strings
- Applications of quantum field theory to condensed matter, cosmology, and particle physics

A. Distributions

Sources and further reading:

- S. D. Joglekar, *Mathematical Physics: Advanced Topics*
- M. Stone and P. Goldbart, *Mathematics for Physics*

A.1. Motivation

Distributions are generalizations of functions. A familiar example of a distribution is the charge density of a point charge which satisfies

$$\rho(\mathbf{x}) = \begin{cases} 0 & \text{if } \mathbf{x} \neq 0 \\ \infty & \text{if } \mathbf{x} = 0 \end{cases} \tag{255}$$

and

$$\int \rho(\mathbf{x}) d^3\mathbf{x} = q\,. \tag{256}$$

The charge density $\rho(\mathbf{x})$ is not a function because it is not well defined at $\mathbf{x} = 0$.

Distributions can also help us make sense of seemingly ill-defined derivatives or integrals of ordinary functions. For example while attempts to evaluate $\frac{d^n}{dx^n}|x|$ at $x = 0$ using the rules of differentiation you learned in first-year calculus will not get you very far, we will see that we can simply and unambiguously evaluate the derivative using a generalization of the derivative appropriate for distributions. Seemingly ambiguous integrals like $\int_{\mathbb{R}} \frac{dx}{x}$

[36] An action has a gauge symmetry if there is a local (as opposed to global) symmetry transformation that leaves it invariant. Despite its name, a gauge symmetry is a redundancy and not a symmetry.

[37] Functional integral quantization is a generalization of path integral quantization. Both functional integral and canonical quantization have their advantages.

[38] A conformal field theory has a scaling symmetry.

[39] Quantum effects that violate classical symmetries

arise in quantum field theory all the time, and one of the ways we can make sense of these integrals is by using distributions.

Limits of sequences of functions may not converge to an ordinary function and some variational problems have no solutions that are ordinary functions. In both cases distributions can help.

A.2. Preliminaries

You already know an example of a distribution that is not a function: the unfortunately named Dirac delta "function." You may have been told that the Dirac delta was the limit of a family of functions

$$\delta(x) \text{“} = \text{”} f(x) = \lim_{\sigma \to 0^+} f_\sigma(x) \tag{257}$$

where the family of functions f_σ satisfies

$$\int_{-\infty}^{\infty} f_\sigma(x) dx = 1 \text{ for all } \sigma > 0 \tag{258}$$

and $f(x) = 0$ for $x \neq 0$. The quotes around the equals sign indicate that this definition is provisional.[40] One such family of functions f_σ is

$$f_\sigma(x) = \begin{cases} \frac{1}{\sigma} & \text{if } |x| < \frac{\sigma}{2} \\ 0 & \text{if } |x| \geq \frac{\sigma}{2} \end{cases}. \tag{259}$$

When do we have to be careful and what can go wrong?

Let's start with some definitions.

Attempt: A function space $\mathcal{F}$ is a vector space whose elements are functions and which has a norm $||\cdot||$ that has the following properties:

- positivity ($||f|| \geq 0$ for all $f \in \mathcal{F}$)
- uniqueness of the identity ($||f|| = 0$ iff $f = 0$)
- triangle inequality ($||f + g|| \leq ||f|| + ||g||$ for all $f, g \in \mathcal{F}$)
- linear homogeneity ($||\lambda f|| = |\lambda| ||f||$ for all $f \in \mathcal{F}$, $\lambda \in \mathbb{C}$)

In an inner product space[41], the norm comes from an inner product. In this case the norm is $||f|| = \sqrt{< f, f >}$. We will mostly be interested in the inner product $< \cdot, \cdot >$ given by

$$< f, g >= \int_a^b dx f^* g \tag{260}$$

[40] There is a way to make this definition make sense.

[41] Unlike the Dirac delta, inner product spaces are fortunately named.

Often we will restrict ourselves to real functions. One important function space is $L^2[a,b]$ which has a norm defined by the above inner product, and has the requirement that its elements are square integrable (i.e. $||f||$ is finite).

But $L^2[a,b]$ does not satisfy our first attempt at a definition of a function space. The zero function clearly has zero norm. So does any function that differs from the zero function at a finite number of points. We need to identify the zero function with these functions if we want a unique identity. This leads us to the following definition:

Definition: A function space $\mathcal{F}$ is a vector space whose elements are *equivalence classes* of functions and which has a norm $\|\cdot\|$ that has the following properties:

- positivity
- uniqueness of the identity
- triangle inequality
- linear homogeneity

Two functions belong to the same equivalence class if they differ at a finite number of points.

The function space $L^2[-\infty,\infty]$ is not general enough to contain the Dirac delta. If we compute the norm of f_σ in (259) we find

$$||f_\sigma||^2 = \int_{-\infty}^{\infty} |f_\sigma(x)|^2 dx \tag{261}$$

$$= \int_{-\sigma/2}^{\sigma/2} \frac{1}{\sigma^2} dx \tag{262}$$

$$= \frac{1}{\sigma} \tag{263}$$

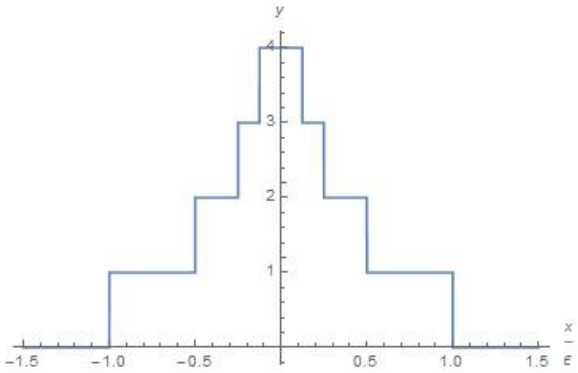
As we take $\sigma \to 0$ the norm blows up, so $||f||$ is undefined and we conclude that δ is not an element of $L^2[-\infty,\infty]$.

We need to be careful if we try to construct δ as a limit of a family of functions. Choose a sequence of functions g_σ [42] that satisfies:

- $\lim_{\sigma\to 0} g_\sigma(x) = 0$ for $x \neq 0$
- $\int_{-\infty}^{\infty} g_\sigma(x)dx = 1$ for all $\sigma > 0$

[42] f_σ from before is one example that satisfies all of the conditions listed.

Figure 1: An infinite tower of rectangles



- For each x and each ϵ we can find σ_0 so that $g_\sigma(x)$ lies inside an infinite tower of rectangles shown in Figure 1 for all $\sigma < \sigma_0$. Each of the rectangles in the infinite tower has height 1. The bottom rectangle has width ϵ and its bottom edge is the x-axis, and each subsequent rectangle has half the width of the previous rectangle. The whole tower has a reflection symmetry about the y-axis. This infinite tower of rectangles gets more and more peaked at the origin as $\epsilon \to 0$ and approaches what we would like to think of as δ.

The limit function $g(x) = \lim_{\sigma\to 0} g_\sigma(x)$ lies inside of the rectangles so

$$\int_{-\infty}^{\infty} g(x)dx < \epsilon + \frac{\epsilon}{2} + \frac{\epsilon}{4} + \cdots \tag{264}$$

$$= 2\epsilon \tag{265}$$

for any ϵ. The order in which we take the limit and perform the integral matters[43]

$$\int_{-\infty}^{\infty} \lim_{\sigma\to 0} g_\sigma(x)dx = 0 \neq 1 = \lim_{\sigma\to 0} \int_{-\infty}^{\infty} g_\sigma(x)dx\,. \tag{266}$$

It is possible to define δ as a limit of functions but we will need to be more careful.

A.3. Test Functions and Distributions

The Dirac delta behaves nicely under integrals. An integral is a map from a function to the real or complex numbers. Distributions map functions to real or complex numbers.

[43] If you are familiar with pointwise versus uniform convergence, the reason the order matters is that the sequence of functions converges to $g(x)$ pointwise but does not converge uniformly. A sequence $f_n(x)$ converges to $f(x)$ pointwise if for each x, $\lim_{n\to\infty} f_n(x) = f(x)$. A sequence $f_n(x)$ converges to $f(x)$ uniformly if for each $\epsilon > 0$ there exists an N such that for all $n > N$ and for all x, $|f_n(x) - f(x)| < \epsilon$.

The space of distributions depends upon which functions we allow our distributions to act. The functions that distributions act upon are known as test functions, and we denote our test function space by $\mathcal{D}$. We sometimes indicate that our distributions map test functions to real numbers using the notation $\mathcal{D}(\mathbb{R})$.

The space of real C^∞ (continuous with an infinite number of derivatives) functions with compact support[44] will be our primary example of a test function space. [45] In general the larger our test function space is, the smaller our space of distributions will be.

We can define a dual vector space $\mathcal{D}^*$ of linear functionals of $\mathcal{D}(\mathbb{R})$ to $\mathbb{R}$. A functional u is an element of $\mathcal{D}^*$ if $u(\phi)$ is a real number for any $\phi \in \mathcal{D}$ and

$$u(a\phi + b\psi) = au(\phi) + bu(\psi) \tag{267}$$

for $a, b \in \mathbb{R}$, and $\phi, \psi \in \mathcal{D}$.

The space of distributions $\mathcal{D}'$ is the largest continuous dual space. In this context continuous means that if a sequence of test functions ϕ_n converges[46] to ϕ in $\mathcal{D}$, then $\lim_{n\to\infty} u(\phi_n) = u(\phi)$ for any u in $\mathcal{D}'$.

We will often use the integral notation $\int \delta(x)\phi(x)dx = \phi(0)$, even though it is slightly imprecise, as it is still a useful way to think about distributions. Our definitions do not give any meaning to the value of a distribution at a point such as $\delta(x)$.

Functions that are[47] locally integrable[48] are distributions.

A.4. Derivation

We cannot define the derivative of a distribution using the ordinary rule

$$u'(x) \stackrel{?}{=} \lim_{\epsilon\to 0} \frac{u(x+\epsilon) - u(x)}{\epsilon} \tag{268}$$

[44] A function with compact support vanishes outside of a finite region.

[45] Much of the following discussion holds if we consider C^∞ functions that decrease faster than any polynomial at $\pm\infty$, or C^n functions with compact support with n sufficiently large instead.

[46] The sequence of test functions ϕ_n converges to ϕ iff the sequence $\phi_n - \phi$ converges to zero in $\mathcal{D}$. A sequence of test functions ψ_n converges to zero if ψ_n and all of its derivatives converge uniformly to zero, and all of the ψ_n vanish outside of the same finite interval $[a, b]$.

[47] To be more precise, locally integrable functions induce distributions. That is for a locally integrable function $f(x)$, we can define an associated distribution f that acts on test functions by $f(\phi) = \int_a^b f(x)\phi(x)dx$. We usually use the same symbol to denote the function and the distribution it induces.

[48] A function is locally integrable if its integral over any compact region is finite.

because $u(x)$ is undefined. For those distributions that are functions we can try another approach

$$u'(\phi) = \int_{-\infty}^{\infty} u'(x)\phi(x)dx \tag{269}$$

$$\stackrel{?}{=} u(x)\phi(x)|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} u(x)\phi'(x)dx \tag{270}$$

$$= -\int_{-\infty}^{\infty} u(x)\phi'(x)dx \tag{271}$$

$$= -u(\phi')\,. \tag{272}$$

We can go from the first line to the second line if $u(x)$ is a function. The next step is justified because our test functions have compact support.

We take this result as a definition of the weak or distributional derivative

$$u'(\phi) \equiv -u(\phi')\,. \tag{273}$$

The n^{th} weak derivative is the obvious generalization

$$u^{(n)}(\phi) \equiv (-1)^n u(\phi^{(n)})\,. \tag{274}$$

Example 1: Heaviside function

The Heaviside function $\Theta(x)$ is given by

$$\Theta(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0\,. \end{cases} \tag{275}$$

The Heaviside function is a distribution because it is manifestly linear and continuous and

$$\Theta(\phi) = \int_{-\infty}^{\infty} \Theta(x)\phi(x)dx \tag{276}$$

$$= \int_{0}^{\infty} \phi(x)dx \tag{277}$$

is a real number since ϕ is finite everywhere and has compact support. (Alternatively we

could note that the Heaviside function is locally integrable.) Let's compute its derivative

$$\Theta'(\phi) = -\Theta(\phi') \tag{278}$$

$$= -\int_0^\infty \phi'(x)dx \tag{279}$$

$$= -(\phi(\infty) - \phi(0)) \tag{280}$$

$$= \phi(0) \tag{281}$$

$$= \delta(\phi)\,. \tag{282}$$

Since ϕ is arbitrary, we have recovered the expected result that $\Theta' = \delta$.

Example 2: $(\ln|x|)'$

Since it is locally integrable, $(\ln|x|)$ is a distribution. Since $\frac{1}{x}$ is not locally integrable, it will be useful to play with our definition of the distribution $(\ln|x|)$ before trying to take the derivative. If we cut out a small region around zero, we do not change the distribution $(\ln|x|)$

$$(\ln|x|)(\phi) \equiv \lim_{\epsilon\to 0^+}\left(\int_{-\infty}^{-\epsilon} + \int_\epsilon^\infty\right)\ln|x|\phi(x)dx\,. \tag{283}$$

The part we cut out vanishes

$$|\lim_{\epsilon\to 0^+}\int_{-\epsilon}^{\epsilon}\ln|x|\phi(x)dx| \le \lim_{\epsilon\to 0^+}\int_{-\epsilon}^{\epsilon}|\ln|x|||\phi(x)|dx \tag{284}$$

$$\le \lim_{\epsilon\to 0^+}\int_{-\epsilon}^{\epsilon}|\ln|x||max(|\phi(x)|)dx \tag{285}$$

$$= max(|\phi(x)|)\lim_{\epsilon\to 0^+} 2\int_0^\epsilon(-\ln x)dx \tag{286}$$

$$= max(|\phi(x)|)\lim_{\epsilon\to 0^+} -2x(\ln x - 1)|_0^\epsilon\,dx \tag{287}$$

$$= 0\,. \tag{288}$$

We can now compute $(\ln|x|)'$

$$(\ln|x|)'(\phi) = -\lim_{\epsilon\to 0^+}\left(\int_{-\infty}^{-\epsilon} + \int_{\epsilon}^{\infty}\right)\ln|x|\phi'(x)dx \tag{289}$$

$$= \lim_{\epsilon\to 0^+}\left[\ln|\epsilon|\,(\phi(\epsilon)-\phi(-\epsilon)) + \left(\int_{-\infty}^{-\epsilon} + \int_{\epsilon}^{\infty}\right)\frac{1}{x}\phi(x)dx\right] \tag{290}$$

$$= \lim_{\epsilon\to 0^+}\left(\int_{-\infty}^{-\epsilon} + \int_{\epsilon}^{\infty}\right)\frac{1}{x}\phi(x)dx \tag{291}$$

$$\equiv \left(PV\frac{1}{x}\right)(\phi)\,. \tag{292}$$

To get the second line we integrated by parts. The third line uses

$$\lim_{\epsilon\to 0^+}\ln|\epsilon|\,(\phi(\epsilon)-\phi(-\epsilon)) = \lim_{\epsilon\to 0^+}\ln|\epsilon|(2\epsilon\phi'(0)) \tag{293}$$

$$= -\lim_{\epsilon\to 0^+} 2\epsilon\phi'(0) \tag{294}$$

$$= 0\,. \tag{295}$$

The distribution defined by the third line is known as the principal value or principal part distribution and after the delta distribution is one of the most important distributions that is not a locally integrable function.

A.5. Multiplication

We can multiply a distribution that is a function by a C^∞ function ψ in the obvious way

$$(\psi u)(\phi) = \int_{-\infty}^{\infty}(\psi(x)u(x))\phi(x)dx \tag{296}$$

$$= \int_{-\infty}^{\infty} u(x)(\psi(x)\phi(x))dx \tag{297}$$

$$= u(\psi\phi)\,. \tag{298}$$

We take this result as a definition for all distributions

$$(\psi u)(\phi) \equiv u(\psi\phi)\,. \tag{299}$$

A.6. Composition

We can try to define composition of a distribution u with a function f in the same way we defined derivation and multiplication. If u is a distribution that is also a function the

natural way to try to define composition $(u \circ f)$ is by making a change of variables $y = f(x)$

$$(u \circ f)(\phi) = \int_{-\infty}^{\infty} u(f(x))\phi(x)dx \tag{300}$$

$$\stackrel{?}{=} \int_{-\infty}^{\infty} u(y)\phi(g(y))|g'(y)|dy \tag{301}$$

where $x = g(y)$. For $(u \circ f)$ to be a distribution, we need $\phi(g(y))|g'(y)|$ to be a test function. It is guaranteed to be a test function if:

- f is C^∞
- $y = f(x)$ has a unique solution $x = g(y)$

In this case we can generalize and define composition via

$$(u \circ f)(\phi) \equiv u((\phi \circ g)|g'|)\,. \tag{302}$$

For specific distributions it is possible to drop some of the above restrictions. We can compose the delta distribution with any function that has only simple roots[49]

$$(\delta \circ f)(\phi) \equiv \sum_i \frac{\delta_{x_i}}{|f'(x_i)|}(\phi) \tag{303}$$

where the sum is over the roots of f. Note that if f has any multiple roots, $(\delta \circ f)$ is undefined.

A.7. Functional Derivation

The solutions to many variational problems are distributions that are not functions. For this reason it can be useful to think about functional derivatives in the language of test functions and distributions. First recall the way you probably learned how to perform functional derivatives:

Example 1: Point particle

Consider a point particle with the action

$$S[q(t), \dot{q}(t)] = \int dt \mathcal{L}(q, \dot{q}) \tag{304}$$

$$= \int dt \left(\frac{1}{2}\dot{q}^2 - V(q)\right). \tag{305}$$

[49] This formula is extremely important!

The variation of the action is

$$\delta S = S[q + \delta q, \dot{q} + \delta\dot{q}] - S[q, \dot{q}] \tag{306}$$

$$= \int dt \delta q \left[\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} \right] + \mathcal{O}(\delta q^2) \,. \tag{307}$$

The term in square brackets is the functional derivative and can be thought of as a generalization of the partial derivative:

$$\delta S = \sum_i \frac{\partial S}{\partial q_i} \delta q_i \tag{308}$$

$$\to \int dt \delta q(t) \frac{\delta S}{\delta q(t)} \tag{309}$$

where instead of a discrete set of q_i we have a continuous set $q(t)$.

The functional derivative can also be defined in terms of test functions if we think of the variation of $q(t)$ as a small parameter ϵ times a test function $\phi(t)$

$$\delta q(t) = \epsilon \phi(t) \tag{310}$$

where depending on the details of the problem it may be necessary to choose a different test function space (for instance to ensure that the test functions vanish at the boundaries of integration).

We can define the functional derivative as

$$\frac{\delta F}{\delta f(x)}[\phi] = \int \frac{\delta F}{\delta f(x)} \phi(x) dx = \lim_{\epsilon \to 0} \frac{F[f + \epsilon\phi] - F[f]}{\epsilon} \,. \tag{311}$$

Example 2: $f^2(x_0)$

Using our new definition we can compute the functional derivative of $F[f] = f^2(x_0)$ with respect to $f(x)$

$$\int \frac{\delta F}{\delta f(x)} \phi(x) dx = \lim_{\epsilon \to 0} \frac{f^2(x_0) + 2\epsilon f(x_0)\phi(x_0) + \epsilon^2 \phi^2(x_0) - f^2(x_0)}{\epsilon} \tag{312}$$

$$= 2 f(x_0) \phi(x_0) \tag{313}$$

$$= \int 2 f(x_0) \delta(x - x_0) \phi(x) dx \tag{314}$$

Since $\phi(x)$ is an arbitrary test function, we have

$$\frac{\delta F}{\delta f(x)} = 2 f(x_0) \delta(x - x_0) \,. \tag{315}$$

The functional derivative has the following properties (which are mostly what you would expect based on the analogy with partial derivatives):

- Linearity:

$$\frac{\delta(\lambda F + \mu G)}{\delta f(x)} = \lambda \frac{\delta F}{\delta f(x)} + \mu \frac{\delta G}{\delta f(x)} \tag{316}$$

- Product Rule:

$$\frac{\delta(FG)}{\delta f(x)} = \frac{\delta F}{\delta f(x)} G + F \frac{\delta G}{\delta f(x)} \tag{317}$$

- Chain Rules:

$$\frac{\delta F[g(f)]}{\delta f(x)} = \frac{\delta F[g(f)]}{\delta g(f(x))} \frac{dg(f(x))}{df(x)} \tag{318}$$

$$\frac{\delta g(F[f])}{\delta f(x)} = \frac{dg(F[f])}{dF[f]} \frac{\delta F[f]}{\delta f(x)} \tag{319}$$

In these equations λ and μ are complex numbers, F and G are functionals, and f and g are functions.

A.8. Green's Functions

We often want to solve differential equations of the form

$$Ly = f \tag{320}$$

where L is a linear differential operator of the form

$$L = p_0(x) \frac{d^n}{dx^n} + p_1(x) \frac{d^{n-1}}{dx^{n-1}} + \cdots + p_n(x) \tag{321}$$

and $y(x)$ and $f(x)$ are functions of a single variable x. The formal solution to (320) is

$$y = L^{-1} f \,. \tag{322}$$

A useful way to think about (322) is by analogy with linear algebra equations on a finite dimensional vector space.

We can think of a function f as an infinite dimensional vector whose elements are the values of the function at each point $f(x)$. In this analogy x labels the components of the vector. A differential operator O is like a matrix, and the associated integral kernel $K(x, y)$ are the elements of the matrix

$$Of(x) = \int K(x, y) f(y) dy \,. \tag{323}$$

The Dirac delta is the analogue of the identity matrix.

There are some additional complications and subtleties when we move from finite-dimensional vector spaces to infinite dimensional vector spaces, but the analogy will be useful to keep in mind for Quantum Field Theory III, where we will encounter more involved expressions involving integral kernels of differential operators.

The inverse operator L^{-1} has an integral kernel $G(x,\xi)$

$$L_x G(x,\xi) = \delta(x-\xi) \tag{324}$$

where the subscript on L_x indicates that the operator acts on the first argument. The integral kernel $G(x,\xi)$ is known as the Green's function.

Because G is the integral kernel of L^{-1}, we have

$$y(x) = \int G(x,\xi) f(\xi) d\xi \tag{325}$$

which satisfies the differential equation

$$\begin{aligned} L_x y(x) &= \int L_x G(x,\xi) f(\xi) d\xi \\ &= \int \delta(x-\xi) f(\xi) d\xi \\ &= f(x)\,. \end{aligned} \tag{326}$$

A.9. Warnings

I will conclude this appendix with a few warnings:

- Distributions act on test functions. You can get nonsensical results if you act with a distribution on something other than a test function, or try to treat a distribution as an ordinary function.
- Multiplication of a distribution with a function only makes sense if that function is a C^∞ function.
- To define composition of a distribution with a function we had to place stringent requirements on the function.
- No mention has been made of products of distributions. The theory of distributions is a linear theory.
- We saw an example where a pointwise limit of sequences of functions $f_n(x) \to f(x)$ behaved very differently from a uniform limit of sequence of functions $f_n \to f$.