Introduction to Quantum Field Theory for Mathematicians

Lecture notes for Math 273, Stanford, Fall 2018

Sourav Chatterjee

(Based on a forthcoming textbook by Michel Talagrand)

Contents

LECTURE 1

Introduction

Date: 9/24/2018
Scribe: Andrea Ottolini

1.1. Preview

This course is intended to be an introduction to quantum field theory for mathematicians. Although quantum mechanics has been successful in explaining many microscopic phenomena which appear to be genuinely random (i.e., the randomness does not stem from the lack of information about initial condition, but it is inherent in the behavior of the particles), it is not a good theory for elementary particles, mainly for two reasons:

- It does not fit well with special relativity, in that the Schrödinger equation is not invariant under Lorentz transformations.
- It does not allow creation or annihilation of particles.

Since in lots of interesting phenomena (e.g., in colliders) particles travel at speeds comparable to the speed of light, and new particles appear after they collide, these aspects have to be taken into account.

Quantum field theory (QFT) is supposed to describe these phenomena well, yet its mathematical foundations are shaky or non-existent. The fundamental objects in quantum field theory are operator-valued distributions. An operator-valued distribution is an abstract object, which when integrated against a test function, yields a linear operator on a Hilbert space instead of a number.

For example, we will define operator-valued distributions a and $a^\dagger$ on $\mathbb{R}^3$ which satisfy that for all $\mathbf{p}, \mathbf{p}' \in \mathbb{R}^3$,

$$[a(\mathbf{p}), a(\mathbf{p}')] = 0, \quad [a^\dagger(\mathbf{p}), a^\dagger(\mathbf{p}')] = 0,$$
$$[a(\mathbf{p}), a^\dagger(\mathbf{p}')] = (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{p}')1,$$

where $[A, B] = AB - BA$ is the commutator, $\delta^{(3)}$ is the Dirac δ on $\mathbb{R}^3$, and 1 denotes the identity operator on an unspecified Hilbert space. For someone with a traditional training in mathematics, it may not be clear what the above statement means. Yet, physics classes on QFT often begin by introducing these operator-valued distributions as if their meaning is self-evident. One of the first objectives of this course is to give rigorous meanings to a and $a^\dagger$, and define the relevant Hilbert space. It turns out

that the correct Hilbert space is the so-called bosonic Fock space, which we will define.

Using a and $a^\dagger$, physicists then define the massive scalar free field φ with mass parameter m, as

$$\varphi(t,\mathbf{x}) = \int_{\mathbb{R}^3} \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\mathbf{p}}}}\Big(e^{-it\omega_{\mathbf{p}}+i\mathbf{x}\cdot\mathbf{p}}a(\mathbf{p}) + e^{it\omega_{\mathbf{p}}-i\mathbf{x}\cdot\mathbf{p}}a^\dagger(\mathbf{p})\Big),$$

where

$$\omega_{\mathbf{p}} = \sqrt{m^2+|\mathbf{p}|^2}.$$

Here $\mathbf{x}\cdot\mathbf{p}$ is the scalar product of $\mathbf{x}$ and $\mathbf{p}$, and $|\mathbf{p}|$ is the Euclidean norm of $\mathbf{p}$. This is an operator-valued distribution defined on spacetime.

Again, it is not at all clear what this means, nor the purpose. We will give a rigorous meaning to all of these and understand where they come from. We will then move on to discuss interacting quantum fields, where the Hilbert space is not clear at all, since the Fock space, which does the job for the free field, is not going to work. Still, computations can be carried out, scattering amplitudes can be obtained, and unrigorous QFT theory leads to remarkably correct predictions for a wide range of phenomena. We will talk about all this and more. In particular, we will talk about φ^4 theory, one-loop renormalization, and the basics of quantum electrodynamics.

1.2. A note on mathematical rigor

Much of quantum field theory is devoid of any rigorous mathematical foundation. Therefore we have no option but to abandon mathematical rigor for large parts of this course. There will be parts where we will not prove theorems with full rigor, but it will be clear that the proofs can be made mathematically complete if one wishes to do so. These are not the problematic parts. However, there will be other parts where no one knows how to put things in a mathematically valid way, and they will appear as flights of fancy to a mathematician. Yet, concrete calculations yielding actual numbers can be carried out in these fanciful settings, and we will go ahead and do so. These situations will be pointed out clearly, and will sometimes be posed as open problems.

1.3. Notation

The following are some basic notations and conventions that we will follow. We will need more notations, which will be introduced in later lectures.

- Throughout these lectures, we will work in units where $\hbar = c = 1$, where $\hbar$ is Planck's constant divided by 2π and c is the speed of light.
- $\mathcal{H}$ is a separable complex Hilbert space.
- If $a \in \mathbb{C}$, a^* is its complex conjugate.

- The inner product of $f, g \in \mathcal{H}$, denoted by (f, g), is assumed to be antilinear in the first variable and linear in the second. In particular, if $\{e_n\}_{n=1}^{\infty}$ is an orthonormal basis of $\mathcal{H}$, and if $f = \sum \alpha_n e_n$ and $g = \sum \beta_n e_n$, then $(f, g) = \sum_{n=1}^{\infty} \alpha_n^* \beta_n$.
- The norm of a state f is denoted by $\|f\|$. A state f is called normalized if $\|f\| = 1$.
- If A is a bounded linear operator on $\mathcal{H}$, $A^\dagger$ denotes its adjoint.
- If A is a bounded linear operator and $A = A^\dagger$, we will say that A is Hermitian. We will later replace this with a more general notion of 'self-adjoint'.
- δ is the Dirac delta at 0, and δ_x is the Dirac delta at x. Among the properties of the delta function, we will be interested in the following two:

$$\int_{-\infty}^{\infty} dz \delta(x - z)\delta(z - y)\xi(z) = \delta(x - y)\xi(z),$$

$$\delta(x) = \lim_{\varepsilon \to 0} \int_{\mathbb{R}} \frac{dy}{2\pi} e^{ixy - \varepsilon y^2} = \frac{1}{2\pi} \int_{\mathbb{R}} dy\, e^{ixy}.$$

- $\widehat{f}(p) = \int_{\mathbb{R}} dx\, e^{-ixp} f(x)$ is the Fourier transform of f.

Note that some of the definitions are slightly different than the usual mathematical conventions, such as that of the Fourier transform. Usually, it is just a difference of sign, but these differences are important to remember.

LECTURE 2

The postulates of quantum mechanics

Date: 9/26/2018
Scribe: Sanchit Chaturvedi

2.1. Postulates 1–4

We will introduce our framework of quantum mechanics through a sequence of five postulates. The first four are given below.

- P1 The state of a physical system is described by a vector in a separable complex Hilbert space $\mathcal{H}$.
- P2 To each (real-valued) observable $\mathcal{O}$ corresponds a Hermitian operator A on $\mathcal{H}$. (It can be a bounded linear operator such that $A = A^\dagger$ but not all A will be like this.)
- P3 If A is the operator for an observable $\mathcal{O}$ then any experimentally observed value of $\mathcal{O}$ must be an eigenvalue of A.
- P4 Suppose that $\mathcal{O}$ is an observable with operator A. Suppose further that A has an orthonormal sequence of eigenvectors $\{x_n\}_{n=1}^{\infty}$ with eigenvalues $\{\lambda_n\}$. Suppose also that the system is in state $\psi \in \mathcal{H}$. Then the probability that the observed value of $\mathcal{O} = \lambda$ is given by

$$\frac{\sum_{i:\lambda_i=\lambda} |(x_i, \psi)|^2}{\|\psi\|^2}.$$

These postulates will be slightly modified later, and replaced with more mathematically satisfactory versions. P5 will be stated later in this lecture.

2.2. A simple example

Consider a particle with two possible spins, say 1 and -1. Then $\mathcal{H} = \mathbb{C}^2$. Consider the observable

$$\mathcal{O} = \begin{cases} +1, & \text{if spin is } 1, \\ -1, & \text{if spin is } -1. \end{cases}$$

Suppose that we take the operator for this observable to be the matrix

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

This has eigenvectors

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

with eigenvalues $1, -1$. If the state of the system is

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \in \mathbb{C}^2,$$

then

$$\text{Prob}(\mathcal{O} = 1) = \frac{|\alpha_1|^2}{|\alpha_1|^2 + |\alpha_2|^2}$$

and

$$\text{Prob}(\mathcal{O} = -1) = \frac{|\alpha_2|^2}{|\alpha_1|^2 + |\alpha_2|^2}.$$

2.3. Adjoints of unbounded operators

DEFINITION 2.1. An unbounded operator A on a Hilbert space $\mathcal{H}$ is a linear map from a dense subspace $\mathcal{D}(A)$ into $\mathcal{H}$.

DEFINITION 2.2. An unbounded operator is called symmetric if

$$(x, Ay) = (Ax, y) \quad \forall x, y \in \mathcal{D}(A).$$

Take any unbounded operator A with domain $\mathcal{D}(A)$. We want to define the adjoint $A^\dagger$. We first define $\mathcal{D}(A^\dagger)$ to be the set of all $y \in \mathcal{H}$ such that

$$\sup_{x \in \mathcal{D}(A)} \frac{|(y, Ax)|}{\|x\|} < \infty.$$

Then for $y \in \mathcal{D}(A^\dagger)$ define $A^\dagger y$ as follows. Define a linear functional $\lambda : \mathcal{D}(A) \to \mathbb{C}$ as

$$\lambda(x) = (y, Ax).$$

Since $y \in \mathcal{D}(A^\dagger)$,

$$c := \sup_{x \in \mathcal{D}(A)} \frac{|(y, Ax)|}{\|x\|} < \infty.$$

Thus $\forall x, x' \in \mathcal{D}(A)$,

$$|\lambda(x) - \lambda(x')| = |(y, A(x - x'))| \leq c\|x - x'\|.$$

This implies that λ extends to a bounded linear functional on $\mathcal{H}$. Hence there exists unique z such that $\lambda(x) = (z, x)$. Let $A^\dagger y := z$.

DEFINITION 2.3. A symmetric unbounded operator is called self-adjoint if $\mathcal{D}(A) = \mathcal{D}(A^\dagger)$, and $A^\dagger = A$ on this subspace.

(In practice we only need to verify $\mathcal{D}(A^\dagger) = \mathcal{D}(A)$, since for any symmetric operator, $\mathcal{D}(A) \subseteq \mathcal{D}(A^\dagger)$, and $A^\dagger = A$ on $\mathcal{D}(A)$.)

DEFINITION 2.4. An operator B is called an extension of A if $\mathcal{D}(A) \subseteq \mathcal{D}(B)$ and $A = B$ on $\mathcal{D}(A)$.

An example is if A is symmetric then $A^\dagger$ is an extension of A.

DEFINITION 2.5. A symmetric operator A is called essentially self-adjoint if it has a unique self-adjoint extension.

2.4. Unitary groups of operators

DEFINITION 2.6. A surjective linear operator $U : \mathcal{H} \to \mathcal{H}$ is called unitary if $\|Ux\| = \|x\|\ \forall x \in \mathcal{H}$.

DEFINITION 2.7. A strongly continuous unitary group $(U(t))_{t\in\mathbb{R}}$ is a collection of unitary operators such that

- $U(s+t) = U(s)U(t)\ \forall s, t \in \mathbb{R}$, and
- for any $x \in \mathcal{H}$ the map $t \mapsto U(t)x$ is continuous.

2.5. Stone's Theorem

There is a one-to-one correspondence between one parameter strongly continuous unitary groups of operators on $\mathcal{H}$ and self-adjoint operators on $\mathcal{H}$. Given U, the corresponding self-adjoint operator A is defined as

$$Ax = \lim_{t\to 0} \frac{U(t)x - x}{it},$$

with $\mathcal{D}(A) = \{x : \text{the above limit exists}\}$. (It is conventional to write $U(t) = e^{itA}$.) Conversely, given any self-adjoint operator A, there is a strongly continuous unitary group $(U(t))_{t\in\mathbb{R}}$ such that the above relation between A and U is satisfied on the domain of A.

2.6. Postulate 5

P5 If the system is not affected by external influences then its state evolves in time as $\psi_t = U(t)\psi$ for some strongly continuous unitary group U that only depends on the system (and not on the state).

By Stone's theorem there exists a unique self-adjoint operator H such that $U(t) = e^{-itH}$. This H is called the 'Hamiltonian'. The Hamiltonian satisfies

$$\begin{aligned}\frac{d}{dt}U(t) &= -iHU(t) = -iHe^{-itH} \\ &= -iU(t)H = -ie^{-itH}H.\end{aligned}$$

We will use the above relations extensively in the sequel.

Besides the five postulates stated above, there is also a sixth postulate about collapse of wavefunctions that we will not discuss (or need) in these lectures.

LECTURE 3

Position and momentum operators

Date: 9/26/2018
Scribe: Sky Cao

3.1. Looking back at Postulate 4

Suppose that A is a self-adjoint operator for an observable $\mathcal{O}$ with an orthonormal sequence of eigenvectors $u_1, u_2, \ldots$ and eigenvalues $\lambda_1, \lambda_2, \ldots$. If the system is in state ψ, Postulate 4 says that the probability that the observed value of $\mathcal{O}$ equals λ is given by

$$\frac{\sum_{i:\lambda_i=\lambda} |(u_i, \psi)|^2}{\|\psi\|^2}.$$

From this, we get the expected value of $\mathcal{O}$:

$$\mathbf{E}_\psi(\mathcal{O}) = \frac{\sum_{i=1}^\infty \lambda_i |(u_i, \psi)|^2}{\|\psi\|^2} = \frac{(\psi, A\psi)}{\|\psi\|^2}.$$

Similarly,

$$\mathbf{E}_\psi(\mathcal{O}^2) = \frac{\sum_{i=1}^\infty \lambda_i^2 |(u_i, \psi)|^2}{\|\psi\|^2} = \frac{(\psi, A^2\psi)}{\|\psi\|^2},$$

and more generally

$$\mathbf{E}_\psi(\mathcal{O}^k) = \frac{(\psi, A^k\psi)}{\|\psi\|^2} \ \forall k.$$

Even more generally, for any α,

$$\mathbf{E}_\psi(e^{i\alpha\mathcal{O}}) = \frac{(\psi, e^{i\alpha A}\psi)}{\|\psi\|^2}.$$

In certain situations, $e^{i\alpha A}$ may be defined by Taylor series expansion. But this is not required; in general we may use Stone's theorem to make sense of $e^{i\alpha A}$.

Now recall that the distribution of a random variable X is completely determined by its characteristic function

$$\phi(\alpha) = \mathbf{E}(e^{i\alpha X}).$$

This allows us to arrive at the following.

Better version of Postulate 4: Suppose that $\mathcal{O}$ is an observable with operator A. If the system is in state ψ, then the characteristic function of this observable at $\alpha \in \mathbb{R}$ is given by

$$\frac{(\psi, e^{i\alpha A}\psi)}{\|\psi\|^2}.$$

Similarly, we have better versions of the other postulates by replacing 'Hermitian' with 'self-adjoint' everywhere.

3.2. A non-relativistic particle in 1-D space

The Hilbert space is $\mathcal{H} = L^2(\mathbb{R})$. The first observable is the position observable. Its operator is denoted X, defined as

$$(X\psi)(x) = x\psi(x).$$

The domain of X is

$$\mathcal{D}(X) = \left\{\psi \in L^2(\mathbb{R}) : \int x^2|\psi(x)|^2 dx < \infty\right\}.$$

It can be verified that X is self-adjoint, and that X does not have an orthonormal sequence of eigenvectors in $L^2(\mathbb{R})$.

In physics, they say that X has 'improper' eigenfunctions. The Dirac δ at x is an improper eigenfunction with eigenvalue x. That is,

$$X\delta_x = x\delta_x.$$

We may verify this in the sense of distributions. For f a test function, we have

$$\int (X\delta_x)(y)f(y)dy = \int y\delta_x(y)f(y)dy = xf(x).$$

On the other hand,

$$\int x\delta_x(y)f(y)dy = xf(x).$$

Thus $X\delta_x = x\delta_x$.

Now suppose the state is ψ. What is the probability distribution of the position of the particle? According to our better version of Postulate 4, the characteristic function of the position variable at $\alpha \in \mathbb{R}$ is

$$\frac{(\psi, e^{i\alpha X}\psi)}{\|\psi\|^2}.$$

It is not hard to show that

$$(e^{i\alpha X}\psi)(x) = e^{i\alpha x}\psi(x),$$

and so the characteristic function is

$$\frac{\int e^{i\alpha x}|\psi(x)|^2 dx}{\int |\psi(x)|^2 dx}.$$

Thus the probability density of the position is

$$\frac{|\psi(x)|^2}{\int |\psi(z)|^2 dz}.$$

The 'position eigenstates' are $\delta_x, x \in \mathbb{R}$. To make this precise, let us approximate $\delta_x(y)$ by

$$\frac{1}{\sqrt{2\pi\varepsilon}} e^{-(y-x)^2/2\varepsilon}$$

as $\varepsilon \to 0$. This state has p.d.f. of the position proportional to $e^{-(y-x)^2/\varepsilon}$. This probability distribution converges to the point mass at x as $\varepsilon \to 0$.

The second observable is the momentum observable. The momentum operator is given by

$$P\psi = -i\frac{d}{dx}\psi,$$

so notationally,

$$P = -i\frac{d}{dx}.$$

We may take the domain of P to be

$$\{\psi \in L^2(\mathbb{R}) : \psi' \in L^2(\mathbb{R})\},$$

but then P will not be self-adjoint. However, one can show that P is essentially self-adjoint, that is, there is a unique extension of P to a larger domain where it is self-adjoint.

Using a similar procedure as with the position, we may show that the probability density of the momentum is given by

$$\frac{|\widehat{\psi}(p)|^2}{\|\widehat{\psi}\|_{L^2}^2},$$

where $\widehat{\psi}$ is the Fourier Transform of ψ. The (improper) momentum eigenstate for momentum p is $\psi(x) = e^{ipx}$, because

$$P\psi = p\psi.$$

Note that ψ is not in $L^2(\mathbb{R})$, but we may approximate

$$\psi(x) \approx e^{ipx - \varepsilon x^2}$$

for ε small. The Fourier Transform of this function is

$$\widehat{\psi}(p') = \int e^{-ip'x} e^{ipx} dx = 2\pi\delta(p - p'),$$

which is proportional to the Dirac delta at p.

LECTURE 4

Time evolution

Date: 10/1/2018
Scribe: Jack Lindsey

4.1. Probability density of momentum

Let us continue our discussion about the 1D non-relativistic particle that we started in the previous lecture. If the system is in state ψ, then we claim that the probability density function of the momentum at $p \in \mathbb{R}$ is

$$\frac{|\widehat{\psi}(p)|^2}{\|\widehat{\psi}\|_{L^2}^2},$$

where $\widehat{\psi}$ is the Fourier transform of ψ. Although a complete proof using our version of Postulate 4 takes some work, it is easy to see why this is true from the following sketch. First, observe that

$$P^k\psi = (-i)^k \frac{d^k}{dx^k}\psi.$$

From this it follows that $\widehat{P^k\psi}(p) = p^k\widehat{\psi}(p)$. On the other hand, by Postulate 4, the expected value of the kth power of the momentum is

$$\frac{(\psi, P^k\psi)}{\|\psi\|_{L^2}^2}.$$

By Parseval's identity, this equals

$$\frac{(\widehat{\psi}, \widehat{P^k\psi})}{\|\widehat{\psi}\|_{L^2}^2} = \int_{-\infty}^{\infty} dp\, p^k \frac{|\widehat{\psi}(p)|^2}{\|\widehat{\psi}\|_{L^2}^2}.$$

This strongly indicates that the p.d.f. of the momentum is proportional to $|\widehat{\psi}(p)|^2$. A complete proof would require that we work with characteristic functions instead of moments.

4.2. The uncertainty principle

Consider the improper state $\psi(x) = \delta_{x_0}(x)$. The position of a particle in this state is fully concentrated at x_0. However, $\widehat{\psi}(p) = e^{-ipx_0}$, which means that the momentum of the particle is 'uniformly distributed on the real line' — which does not make mathematical sense, but can be thought of as an idealization of a very spread out probability distribution. On the other

hand, if $\psi(x) = e^{ip_0 x}$, then the position variable has a similar uniform spread on the real line, but $\widehat{\psi}(p) = \delta(p - p_0)$, which means that the momentum is fully concentrated at p_0.

The above discussion indicates that it may be impossible to have a state that localizes both the position and the momentum. Indeed, this is a result that can be proved in great generality, as follows. Consider a general quantum system. Given an observable $\mathcal{O}$ with operator A, it follows from Postulate 4 that if the system is in state ψ, then the standard deviation of $\mathcal{O}$ is given by

$$\sqrt{\frac{(\psi, A^2\psi)}{\|\psi\|^2} - \frac{(\psi, A\psi)^2}{\|\psi\|^4}}.$$

Let us denote this standard deviation by $\Delta_\psi(A)$. The following result is known as the Heisenberg Uncertainty Principle (rigorous statement and proof due to H. Weyl):

THEOREM 4.1. *For any self-adjoint A and B and any $\psi \in \mathcal{D}(A) \cap \mathcal{D}(B)$ such that $A\psi \in \mathcal{D}(B)$ and $B\psi \in \mathcal{D}(A)$, we have*

$$\Delta_\psi(A)\Delta_\psi(B) \geq \frac{1}{4}|(\psi, [A, B]\psi)|^2$$

where $[A, B] = AB - BA$.

PROOF SKETCH. Apply the Cauchy–Schwarz inequality, starting from the right-hand side and expanding. □

4.3. The uncertainty principle for position and momentum

Now let us see what the uncertainty principle gives for the position and momentum operators in 1D. For any ψ, we have

$$(XP\psi)(x) = X(-i\psi')(x) = -ix\psi'(x).$$

On the other hand,

$$(PX\psi)(x) = -i\frac{d}{dx}(x\psi(x)) = -ix\psi'(x) - i\psi(x).$$

Putting these together, we see that $XP - PX = i \cdot 1$, where 1 is the identity operator. Hence, by the uncertainty principle,

$$\Delta_\psi(X)\Delta_\psi(P) \geq \frac{1}{4}$$

for all normalized ψ for which $[X, P]\psi$ is defined and in L^2, in units where $\hbar = 1$. For example, this holds for any Schwartz function ψ with L^2 norm equal to one. Note that this implies that position and momentum cannot be simultaneously localized — the p.d.f. of at least one must have substantial standard deviation.

4.4. Free evolution in 1D

The evolution of the state of a particle in 1D in the absence of a potential (i.e., free evolution) is governed by the Hamiltonian

$$H = -\frac{1}{2m}\frac{d^2}{dx^2},$$

where m is the mass of the particle. (This is a physical fact, with no mathematical justification.) The operator H is essentially self-adjoint, and we will identify it with its unique self-adjoint extension.

If ψ is the state at time 0 and ψ_t is the state at time t, then $\psi_t = U(t)\psi = e^{-itH}\psi$. This implies the Schrödinger equation:

$$\frac{\partial \psi_t}{\partial t} = -iH\psi_t = \frac{i}{2m}\frac{\partial^2 \psi_t}{\partial x^2}.$$

(Again, recall that we are working in units where $\hbar = 1$.) This equation has an explicit solution, which we will not write down. It is easy to show using the explicit form of the solution that for any $\psi \in L^2(\mathbb{R})$, $\|\psi_t\|_{L^\infty} \to 0$ as $t \to \infty$. So the p.d.f. of the position flattens as $t \to \infty$. In other words, there is no limiting probability distribution for the position of the particle; it spreads out across the whole real line.

4.5. Free evolution of the momentum

Now let us see what happens to the momentum of a free particle in 1D. It is not hard to show, using the Schrödinger equation, that

$$\widehat{\psi}_t(p) = e^{-itp^2/2m}\widehat{\psi}(p).$$

Thus, $|\widehat{\psi}_t(p)|^2 = |\widehat{\psi}(p)|^2$. Hence the p.d.f. of the momentum is not changing at all. This is what we expect, since the momentum of a free particle should not be changing (although, of course, the randomness of the result of a given observation is unexpected from a classical perspective).

4.6. A particle in a potential

Now let us briefly consider particles in the presence of a potential V. The Hamiltonian is now given by

$$H = -\frac{1}{2m}\frac{d^2}{dx^2} + V,$$

where V as an operator is defined as $(V\psi)(x) = V(x)\psi(x)$.

When is H is essentially self-adjoint? The answer to this question is not at all obvious. There are many results in the literature, quite comprehensively surveyed in Reed and Simon. For example, one useful result is the following.

THEOREM 4.2. *The Hamiltonian H is essentially self-adjoint if $V = V_1 + V_2$ for some $V_1 \in L^2(\mathbb{R})$ and $V_2 \in L^\infty(\mathbb{R})$.*

This result does not work for growing potentials. For instance, it does not cover the case of a simple harmonic oscillator, where $V(x)$ grows quadratically in x. For such cases, the following result, due to Kato, is helpful.

THEOREM 4.3. *The Hamiltonian H is essentially self-adjoint if V is locally L^2 and $V(x) \geq -V^*(|x|)$ for some V^* such that $V^*(r) = o(r^2)$ as $r \to \infty$.*

(Note that in particular if $V(x)$ is locally L^2 and lower-bounded by a constant, then it satisfies the condition.)

4.7. Simple harmonic oscillator

Theorem 4.3 says, for example, that H is essentially self-adjoint if

$$V(x) = \frac{1}{2} m\omega^2 x^2.$$

This is the potential for a simple harmonic oscillator with frequency ω. Moreover, the corresponding Hamiltonian

$$H = -\frac{1}{2m}\frac{d^2}{dx^2} + \frac{1}{2} m\omega^2 x^2$$

has a complete orthonormal sequence of eigenvectors. For simplicity take $m = \omega = 1$. Then the (orthonormal) eigenvectors are

$$e_n(x) = C_n H_n(x) e^{-x^2/2},$$

where C_n is the normalization constant, and

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n}(e^{-x^2})$$

is the 'nth physicist's Hermite polynomial'.

4.8. Bound states

Note that if H is a Hamiltonian and ψ is an eigenfunction with eigenvalue λ, then the evolution of this eigenfunction under this Hamiltonian is given by $\psi_t = e^{-itH}\psi = e^{-it\lambda}\psi$. So the p.d.f. of the position does not change over time.

Physicists call this a 'bound state'. This means that if you are in the state, you will not freely move out of it (e.g., the p.d.f. of the position will not become more and more flat). If you have a potential which allows particles to move out freely, then the Hamiltonian cannot have a complete orthonormal sequence of eigenstates as in the previous example.

4.9. Dirac notation

Vectors in a Hilbert space are denoted by $|x\rangle$ (these care called 'ket vectors'). Often, we will write vectors like $|0\rangle, |1\rangle, |2\rangle$, etc. Physicists will say, for example, that $|\lambda\rangle$ is an eigenvector with eigenvalue λ. Just like mathematicians would write $x_1, x_2, x_3, \ldots$, physicists write $|1\rangle, |2\rangle, |3\rangle, \ldots$.

We also have 'bra vectors': $\langle x|$ is the linear functional taking $y \mapsto (x, y)$. With this notation, we have

- $\langle x|y\rangle$ is the action of x on y, equal to (x, y).
- $\langle \alpha x|y\rangle = \alpha^* \langle x|y\rangle$ and $\langle x|\alpha y\rangle = \alpha \langle x|y\rangle$.
- (x, Ay) is written as $\langle x|A|y\rangle$. Note that $A|y\rangle = Ay$.

One of the great uses of this notation is the following. Let $|1\rangle, |2\rangle, |3\rangle, \ldots$ be an orthonormal basis of $\mathcal{H}$. Then $\sum_{i=1}^{\infty} |i\rangle \langle i| = 1$, meaning that

$$\left(\sum_{i=1}^{\infty} |i\rangle \langle i|\right) |x\rangle = \sum_{i=1}^{\infty} |i\rangle \langle i|x\rangle = |x\rangle .$$

This is very useful; often one replaces 1 with such an expression in a derivation. Going even further, on $L^2(\mathbb{R})$, a physicist would write that

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} dp\, |p\rangle \langle p| = 1,$$

where $|p\rangle (x) = e^{ipx}$, which is an improper element of L^2. The derivation is as follows:

$$\begin{aligned}\left(\frac{1}{2\pi} \int_{-\infty}^{\infty} dp\, |p\rangle \langle p|\right)\psi(x) &= \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} dp\, |p\rangle \langle p|\psi\rangle\right)(x) \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dp \widehat{\psi}(p) e^{ipx} = \psi(x).\end{aligned}$$

LECTURE 5

Many particle states

Date: 10/3/2018
Scribe: Philip Ken-Ka Hwang

5.1. Position and momentum space wavefunctions

Consider a particle with state space $\mathcal{H} = L^2(\mathbb{R})$. Until now, we have represented a state $|\psi\rangle \in \mathcal{H}$ as a function $\psi \in L^2(\mathbb{R})$. This is called the position space representation of ψ. We can also identify $|\psi\rangle$ with $\widehat{\psi}$, the Fourier transform of ψ. The function $\widehat{\psi}$ is called the momentum space representation of $|\psi\rangle$, or simply the momentum wavefunction. Physicists view ψ and $\widehat{\psi}$ as different representations of the same abstract object $|\psi\rangle$.

Example 5.1. Consider the improper momentum eigenstate $|p_0\rangle$ for $p_0 \in \mathbb{R}$. The position space representation is the function $\psi(x) = e^{ip_0 x}$. The Fourier transform of this function is $\widehat{\psi}(p) = 2\pi\delta(p - p_0)$. The function $\widehat{\psi}$ is the momentum space representation of the state $|p_0\rangle$. Often $\widehat{\psi}$ is also written as ψ if it is clear from the context that we are working in momentum space.

We have seen that if ψ is the position space representative of a free particle, then ψ evolves as $\psi_t = e^{-itH}\psi$, where

$$H = -\frac{1}{2m}\frac{d^2}{dx^2}.$$

The momentum space representative $\varphi = \widehat{\psi}$ evolves as $\varphi_t(p) = e^{-itp^2/2m}\varphi(p)$. Furthermore the Hamiltonian on the momentum space is

$$H\varphi(p) = \frac{p^2}{2m}\varphi(p).$$

If we distinguish the two Hamiltonian operators as H_{x} and H_{p}, then they are related as

$$\widehat{H_{\mathrm{x}}\psi} = H_{\mathrm{p}}\widehat{\psi}.$$

Usually, such a distinction is not made.

5.2. Schrödinger and Heisenberg pictures

Let ψ be a normalized state. The state ψ evolves in times as $\psi_t = e^{-itH}$. If you have an observable with operator A, then its expected value at time t is $\langle\psi_t|A|\psi_t\rangle = (\psi_t, A\psi_t)$. This is called the Schrödinger picture of time

evolution of states. The Heisenberg picture is a different way of looking at time evolution, where operators evolve instead of states. If A is the operator for an observable, then in the Heisenberg picture, the operator for the same observable at time t is $A_t = e^{itH}Ae^{-itH}$. The expected value of the observable at time t is just the expected value of A_t computed at time 0, that is, $\langle\psi|A_t|\psi\rangle$. Note that this gives the same result as in the Schrödinger picture, since

$$\langle\psi|A_t|\psi\rangle = \langle\psi|e^{itH}Ae^{-itH}|\psi\rangle = \langle\psi_t|A|\psi_t\rangle.$$

Thus, the two pictures are different ways of carrying out the same calculations. In quantum field theory, however, it often helps to think in terms of the Heisenberg picture.

5.3. Tensor product state spaces

Let $\mathcal{H}$ be the Hilbert space for one particle. Then the Hilbert space for n such particles is the n-fold tensor product $\mathcal{H}^{\otimes n}$. The tensor product is defined as follows.

If we have $\psi_1, \psi_2, \ldots, \psi_n \in H$, we want to define a tensor product $\psi_1 \otimes \psi_2 \otimes \cdots \otimes \psi_n$ which is multilinear, that is,

(1) $\psi_1 \otimes \cdots \otimes (\alpha\psi_i) \otimes \cdots \otimes \psi_n = \alpha(\psi_1 \otimes \cdots \otimes \psi_n)$, and
(2) $\psi_1 \otimes \cdots \otimes (\psi_i + \psi_i') \otimes \cdots \otimes \psi_n = \psi_1 \otimes \cdots \otimes \psi_i \otimes \cdots \otimes \psi_n + \psi_1 \otimes \cdots \otimes \psi_i' \otimes \cdots \otimes \psi_n$.

To define this product, let us start with an orthonormal basis $e_1, e_2, \ldots$ of $\mathcal{H}$. Consider all formal linear combinations like

$$\sum_{i_1,\ldots,i_n \geq 1} \alpha_{i_1\cdots i_n} e_{i_1} \otimes \cdots \otimes e_{i_n}$$

such that $\sum |\alpha_{i_1\cdots i_n}|^2 < \infty$. On this set, we can define addition and scalar multiplication in the natural way. We can also define an inner product:

$$\left(\sum \alpha_{i_1\cdots i_n} e_{i_1} \otimes \cdots \otimes e_{i_n}, \sum \beta_{i_1\cdots i_n} e_{i_1} \otimes \cdots \otimes e_{i_n}\right)$$
$$:= \sum \alpha^*_{i_1\cdots i_n} \beta_{i_1\cdots i_n}.$$

One can check that this is a Hilbert space. Since this is a basis-dependent construction, call this space $\mathcal{H}_e^{\otimes n}$. Suppose that $f_1, f_2, \ldots$ is another orthonormal basis. Say $e_i = \sum_{j=1}^{\infty} a_{ij} f_j$. We can then produce a natural isomorphism $\psi : \mathcal{H}_e^{\otimes n} \to \mathcal{H}_f^{\otimes n}$ as

$$\psi(e_{i_1} \otimes \cdots \otimes e_{i_n}) = \sum_{j_1,\ldots,j_n} a_{i_1 j_1} a_{i_2 j_2} \cdots a_{i_n j_n} f_{j_1} \otimes \cdots \otimes f_{j_n}.$$

Such isomorphisms allow us to identify the various constructions for various choices of bases, and define a single abstract object called $\mathcal{H}^{\otimes n}$.

What is $\psi_1 \otimes \cdots \otimes \psi_n$ for $\psi_1, \ldots, \psi_n \in \mathcal{H}$? Take any basis $e_1, e_2, \ldots$. Suppose that $\psi_i = \sum_{j=1}^{\infty} a_{ij} e_j$. Define

$$\psi_1 \otimes \cdots \otimes \psi_n := \sum_{j_1, \ldots, j_n} a_{1j_1} a_{2j_2} \ldots a_{nj_n} e_{j_1} \otimes \cdots \otimes e_{j_n}.$$

This is a basis-dependent map from $\mathcal{H}^n$ into $\mathcal{H}_e^{\otimes n}$. However, we can have a basis-independent definition of $\psi_1 \otimes \cdots \otimes \psi_n$ by observing that the diagram

$$\begin{array}{ccc} \mathcal{H}^n & \longrightarrow & \mathcal{H}_e^{\otimes n} \\ & \searrow & \uparrow \\ & & \mathcal{H}_f^{\otimes n} \end{array}$$

commutes. This implies that $(\psi_1, \ldots, \psi_n) \mapsto \psi_1 \otimes \cdots \otimes \psi_n \in \mathcal{H}^{\otimes n}$ is well-defined.

EXAMPLE 5.2. Suppose $\mathcal{H} = L^2(\mathbb{R})$ and $e_1, e_2, \ldots$ is an orthonormal basis. Then an element of $\mathcal{H}_e^{\otimes n}$ is of the form $\sum \alpha_{i_1 \cdots i_n} e_{i_1} \otimes \cdots \otimes e_{i_n}$. We can map this element into $L^2(\mathbb{R}^n)$ as

$$\psi(x_1, \ldots, x_n) = \sum \alpha_{i_1 \cdots i_n} e_{i_1}(x_1) e_{i_2}(x_2) \cdots e_{i_n}(x_n).$$

It is straightforward to check that this map is an isomorphism. If we use a different orthonormal basis $f_1, f_2, \ldots$, then the isomorphic image of this element in $H_f^{\otimes n}$ also maps to the same function $\psi \in L^2(\mathbb{R}^n)$. If $\psi_1, \ldots, \psi_n \in L^2(\mathbb{R})$ and $\psi = \psi_1 \otimes \cdots \otimes \psi_n$, then ψ, as an element of $L^2(\mathbb{R}^n)$, is given by $\psi(x_1, \ldots, x_n) = \psi_1(x_1) \cdots \psi_n(x_n)$

5.4. Time evolution on a tensor product space

Suppose that the state of a single particle evolves according to the unitary group $(U(t))_{t \in \mathbb{R}}$. Then the time evolution on $\mathcal{H}^{\otimes n}$ of n non-interacting particles, also denoted by $U(t)$, is defined as

$$U(t)(\psi_1 \otimes \cdots \otimes \psi_n) := (U(t)\psi_1) \otimes (U(t)\psi_2) \otimes \cdots \otimes (U(t)\psi_n)$$

and extended by linearity. (It is easy to check that this is well-defined.) Consequently, the Hamiltonian is given by

$$\begin{aligned} H(\psi_1 \otimes \cdots \otimes \psi_n) &= -\lim_{t \to 0} \frac{1}{it} (U(t)(\psi_1 \otimes \cdots \otimes \psi_n) - \psi_1 \otimes \cdots \otimes \psi_n) \\ &= -\lim_{t \to 0} \frac{1}{it} (U(t)\psi_1 \otimes U(t)\psi_2 \otimes \cdots \otimes U(t)\psi_n - \psi_1 \otimes \psi_2 \otimes \cdots \otimes \psi_n) \\ &= -\lim_{t \to 0} \frac{1}{it} \sum_{j=1}^{n} U(t)\psi_1 \otimes \cdots \otimes U(t)\psi_{j-1} \otimes (U(t)\psi_j - \psi_j) \otimes \psi_{j+1} \otimes \cdots \otimes \psi_n \\ &= \sum_{j=1}^{n} \psi_1 \otimes \cdots \otimes \psi_{j-1} \otimes H\psi_j \otimes \psi_{j+1} \otimes \cdots \otimes \psi_n. \end{aligned}$$

5.5. Example of time evolution on a product space

Take $\mathcal{H} = L^2(\mathbb{R})$ and let $U(t)$ be the free evolution group, generated by the Hamiltonian

$$H = -\frac{1}{2m}\frac{d^2}{dx^2}.$$

From the previous lecture we know that $\mathcal{H}^{\otimes n} = L^2(\mathbb{R}^n)$. Moreover if $\psi = \psi_1 \otimes \cdots \otimes \psi_n$, then as a function, $\psi(x_1, \ldots, x_n) = \psi_1(x_1) \cdots \psi_n(x_n)$. Therefore,

$$\begin{aligned} H\psi &= \sum_{i=1}^{n} \psi_1(x_1) \cdots \psi_{i-1}(x_{i-1}) \left(-\frac{1}{2m} \psi_i''(x_i) \right) \psi_{i+1}(x_{i+1}) \ldots \psi_n(x_n) \\ &= -\frac{1}{2m} \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2} (\psi_1(x_1) \ldots \psi_n(x_n)) \\ &= -\frac{1}{2m} \Delta\psi. \end{aligned}$$

So by linearity,

$$H\psi = -\frac{1}{2m}\Delta\psi$$

for each $\psi \in \mathcal{D}(\Delta)$, where $\mathcal{D}(\Delta)$ is the domain of the unique self-adjoint extension of the Laplacian.

LECTURE 6

Bosonic Fock space

Date: 10/5/2018
Scribe: Anav Sood

6.1. Bosons

There are two kinds of elementary particles — bosons and fermions. We will deal only with bosons for now. Let $\mathcal{H}$ be the single particle state space for any given particle that is classified as a boson. The main postulate about bosons is that the state of a system of n such particles is always a member of a certain subspace of $\mathcal{H}^{\otimes n}$, which we denote by $\mathcal{H}_{sym}^{\otimes n}$ and define below.

Let $e_1, e_2, \ldots$ be an orthonormal basis of $\mathcal{H}$. Define $\mathcal{H}_{e,sym}^{\otimes n}$ to be all elements of the form $\sum \alpha_{i_1 \cdots i_n} e_{i_1} \otimes \cdots \otimes e_{i_n}$ such that

$$\alpha_{i_1 \cdots i_n} = \alpha_{i_{\sigma(1)} \cdots i_{\sigma(n)}}$$

for all $i_1, \ldots, i_n$ and $\sigma \in S_n$, where S_n is the group of all permutations of $1, \ldots, n$. It turns out that this is a basis-independent definition, in the sense that the natural isomorphism between $\mathcal{H}_e^{\otimes n}$ and $\mathcal{H}_f^{\otimes n}$ discussed earlier is also an isomorphism between the corresponding $\mathcal{H}_{e,sym}^{\otimes n}$ and $\mathcal{H}_{f,sym}^{\otimes n}$. Thus we can simply refer to $\mathcal{H}_{sym}^{\otimes n}$. Moreover, this is a closed subspace and hence a Hilbert space.

For example, take $\mathcal{H} = L^2(\mathbb{R})$ so $\mathcal{H}^{\otimes n} = L^2(\mathbb{R}^n)$. Then it is not hard to show that

$$\mathcal{H}_{sym}^{\otimes n} = \{\psi \in L^2(\mathbb{R}^n) : \psi(x_1 \ldots, x_n) = \psi(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \text{ for all } \sigma \in S_n\}.$$

Another important fact is that if $U(t)$ is any evolution on $\mathcal{H}$, then its extension to $\mathcal{H}^{\otimes n}$ maps $\mathcal{H}_{sym}^{\otimes n}$ into itself.

Next, let us consider the problem of finding an orthonormal basis for $\mathcal{H}_{sym}^{\otimes n}$, starting with an orthonormal basis $e_1, e_2, \ldots$ of $\mathcal{H}$. Take $m_1, \ldots, m_n$ and consider the vector

$$\sum_{\sigma \in S_n} e_{m_{\sigma(1)}} \otimes \cdots \otimes e_{m_{\sigma(n)}}.$$

This element does not have norm 1, so in attempt to construct an orthonormal basis we will normalize it. First, for each $i \geq 1$ let

$$n_i = |\{j : m_j = i, 1 \leq j \leq n\}|.$$

Note that $\sum_{i=1}^{\infty} n_i = n$. For example, if $n = 3$, $m_1 = 1$, $m_2 = 1$ and $m_3 = 2$, then the element in question is

$$e_1 \otimes e_1 \otimes e_2 + e_1 \otimes e_1 \otimes e_2 + e_1 \otimes e_2 \otimes e_1 + e_1 \otimes e_2 \otimes e_1 + e_2 \otimes e_1 \otimes e_1 + e_2 \otimes e_1 \otimes e_1,$$

and $n_1 = 2$, $n_2 = 1$, $n_3 = n_4 = \cdots = 0$.

It is an easy combinatorial exercise to show that

$$\left\| \sum_{\sigma \in S_n} e_{m_{\sigma(1)}} \otimes \cdots \otimes e_{m_{\sigma(n)}} \right\|^2 = n! \prod_{i=1}^{\infty} n_i!.$$

Therefore the element

$$\frac{1}{\sqrt{n! \prod_{i=1}^{\infty} n_i!}} \sum_{\sigma \in S_n} e_{m_{\sigma(1)}} \otimes \cdots \otimes e_{m_{\sigma(n)}}$$

has norm one. Additionally it is completely determined by the numbers $n_1, n_2 \ldots$, so we can choose to meaningfully denote it by $|n_1, n_2, \ldots\rangle$, which is the usual convention. Note that this notation is dependent on the basis $e_1, e_2, \ldots$ of $\mathcal{H}$. It is not hard to show that

$$\langle n_1', n_2', \ldots | n_1, n_2, \ldots \rangle = \delta_{n_1 n_1'} \delta_{n_2 n_2'} \cdots,$$

where $\delta_{k,l}$ is the Kronecker delta function:

$$\delta_{k,l} = \begin{cases} 1 & \text{if } k = l, \\ 0 & \text{if } k \neq l. \end{cases}$$

In fact, more is true: The set of elements

$$\left\{ |n_1, n_2, \ldots\rangle : n_1, n_2, \ldots \geq 0, \sum_{i=1}^{\infty} n_i = n \right\}$$

form an orthonormal basis of $\mathcal{H}_{sym}^{\otimes n}$. Note that this basis for $\mathcal{H}_{sym}^{\otimes n}$ depends on the initial choice of basis $e_1, e_2, \ldots$ for $\mathcal{H}$.

Lastly, we define $\mathcal{H}_{sym}^{\otimes 0} := \mathbb{C}$ with basis $\{1\}$, The basis element 1 is denoted by $|0, 0, 0, \ldots\rangle$ or simply just $|0\rangle$. This is called the vacuum state. It is important to note that $|0\rangle \neq 0$.

6.2. Bosonic Fock Space

Consider a Hilbert space $\mathcal{H}$. We define $\mathcal{B}$ to be the set of sequences $(\psi_0, \psi_1, \ldots)$ where $\psi_n \in \mathcal{H}_{sym}^{\otimes n}$ for all n and $\sum_{n=0}^{\infty} \|\psi_n\|^2 < \infty$. We will formally notate this sequence as $\sum_{n=0}^{\infty} \psi_n$. We can put an inner product on $\mathcal{B}$ as

$$(\psi, \phi) = \sum_{i=0}^{\infty} (\psi_n, \phi_n).$$

This makes $\mathcal{B}$ a Hilbert space. This is called the Bosonic Fock space of $\mathcal{H}$. We can interpret a state $\psi \in \mathcal{B}$ in the following way. The system

has n particles with probability $\|\psi_n\|^2 / \sum_{i=1}^{\infty} \|\psi_i\|^2$, and given that it has n particles, the state of the system is ψ_n.

Let $e_1, e_2, \ldots$ be an orthonormal basis of $\mathcal{H}$. Then a straightforward argument shows that the set

$$\left\{ |n_1, n_2, \ldots \rangle : n_i \geq 0 \text{ for all } i, \sum_{i=1}^{\infty} n_i < \infty \right\}$$

is an orthonormal basis for $\mathcal{B}$. Note this basis depends on the basis chosen for $\mathcal{H}$. Next let

$$\mathcal{B}_0 := \bigoplus_{n=0}^{\infty} \mathcal{H}_{sym}^{\otimes n} = \left\{ \sum_{n=0}^{\infty} \psi_n \in \mathcal{B} : \psi_n = 0 \text{ for all but finitely many } n \right\}.$$

This is a dense subspace of $\mathcal{B}$. All our operators on $\mathcal{B}$ will be defined on this dense subspace.

The process of constructing a Fock space, starting from a single particle Hilbert space, is known as **second quantization**.

LECTURE 7

Creation and annihilation operators

Date: 10/8/2018
Scribe: Casey Chu

7.1. Operator-valued distributions

Fix a basis of $\mathcal{H}$ and for each $k \geq 1$, let $a_k^\dagger : \mathcal{B}_0 \to \mathcal{B}$ be a linear operator defined on basis elements as

$$a_k^\dagger |n_1, n_2, \dots\rangle = \sqrt{n_k + 1} |n_1, n_2, \dots, n_{k-1}, n_k + 1, n_{k+1}, \dots\rangle.$$

Note that $a_k^\dagger$ maps $\mathcal{H}_{sym}^{\otimes n}$ into $\mathcal{H}_{sym}^{\otimes (n+1)}$. Thus it 'creates' a particle, and is therefore called a 'creation operator'. Note, in particular, that

$$\begin{aligned} a_k^\dagger |0\rangle &= a_k^\dagger |0, 0, \dots\rangle \\ &= |\underbrace{0, \dots, 0}_{k-1 \ 0s}, 1, \dots\rangle. \end{aligned}$$

Next, for all k, let $a_k : \mathcal{B}_0 \to \mathcal{B}$ be a linear operator defined on basis elements as

$$a_k |n_1, n_2, \dots\rangle = \begin{cases} \sqrt{n_k} |n_1, n_2, \dots, n_{k-1}, n_k - 1, n_{k+1}, \dots\rangle & \text{if } n_k \geq 1, \\ 0 & \text{if } n_k = 0. \end{cases}$$

Again, note that a_k maps $\mathcal{H}_{sym}^{\otimes n}$ into $\mathcal{H}_{sym}^{\otimes (n-1)}$ for $n \geq 1$. Thus it 'destroys' a particle, and is therefore called an 'annihilation operator'. Note, in particular, that

$$a_k |\underbrace{0, \dots, 0}_{k-1 \ 0s}, 1, \dots\rangle = 1 \cdot |0, 0, \dots\rangle = |0\rangle.$$

Now using these operators we will define operator-valued distributions on $\mathcal{H}$, namely an object which takes in a function as input and returns an operator as output. Take any $f \in \mathcal{H}$ and let $\sum_{k=1}^\infty \alpha_k e_k$ be its expansion in the chosen basis. Define

$$A(f) = \sum_{k=1}^\infty \alpha_k^* a_k, \quad A^\dagger(f) = \sum_{k=1}^\infty \alpha_k a_k^\dagger.$$

Note that A and $A^\dagger$ map from $\mathcal{H}$ into the set of linear operators from $\mathcal{B}_0$ into $\mathcal{B}$. Although these are defined in a basis-dependent way, we will now show that A and $A^\dagger$ are actually basis-independent.

7.2. Basis-independence of A

First, we will show that A is basis-independent. We will show this only for $\mathcal{H} = L^2(\mathbb{R})$, but the proof extends to general $\mathcal{H}$ using isometries between Hilbert spaces.

Let us start with A. Fix an orthonormal basis $e_1, e_2, \ldots \in \mathcal{H}$. Consider a basis element $|n_1, n_2, \ldots\rangle$ of $\mathcal{B}$, where $\sum_{i=1}^{\infty} n_i = n$ and hence $|n_1, n_2, \ldots\rangle \in \mathcal{H}_{sym}^{\otimes n} = L^2(\mathbb{R}^n)_{sym}$. Let φ denote this function, which is a symmetric function in n variables. More explicitly, choose a list of integers $m_1, m_2, \ldots, m_n$ such that for each i, n_i counts the number of i's listed. Then

$$\varphi(x_1, \ldots, x_n) = \frac{1}{\sqrt{n! \prod_i n_i!}} \sum_{\sigma \in S_n} e_{m_{\sigma(1)}}(x_1) \cdots e_{m_{\sigma(n)}}(x_n).$$

Let $\psi = a_k \varphi$. We know that

$$\psi = \sqrt{n_k}\, |n_1, n_2, \ldots, n_k - 1, ...\rangle \in L^2(\mathbb{R}^{n-1})_{sym},$$

so we may similarly obtain the following explicit expression for ψ. Let $m'_1, m'_2, \ldots, m'_{n-1}$ be integers obtained by removing one k from $m_1, \ldots, m_n$. Then

$$\begin{aligned} &\psi(x_1, \ldots, x_{n-1}) \\ &= \frac{\sqrt{n_k}}{\sqrt{(n-1)!(n_k - 1)! \prod_{i \neq k} n_i!}} \sum_{\sigma \in S_{n-1}} e_{m'_{\sigma(1)}}(x_1) \cdots e_{m'_{\sigma(n-1)}}(x_{n-1}). \end{aligned}$$

Now note that

$$\begin{aligned} &\int_{-\infty}^{\infty} e_k(y)^* \sum_{\sigma \in S_n} e_{m_{\sigma(1)}}(x_1) \cdots e_{m_{\sigma(n-1)}}(x_{n-1})\, e_{m_{\sigma(n)}}(y)\, dy \\ &\quad = \sum_{\sigma \in S_n} e_{m_{\sigma(1)}}(x_1) \cdots e_{m_{\sigma(n-1)}}(x_{n-1}) \int_{-\infty}^{\infty} e_k(y)^* e_{m_{\sigma(n)}}(y)\, dy \\ &\quad = \sum_{\substack{\sigma \in S_n \\ m_{\sigma(n)} = k}} e_{m_{\sigma(1)}}(x_1) \cdots e_{m_{\sigma(n-1)}}(x_{n-1}) \\ &\quad = n_k \sum_{\sigma \in S_{n-1}} e_{m_{\sigma(1)}}(x_1) \cdots e_{m_{\sigma(n-1)}}(x_{n-1}), \end{aligned}$$

using the fact that

$$\int_{-\infty}^{\infty} e_j(y)^* e_k(y)\, dy = \delta_{jk}$$

to go from the second line to the third. We recognize that the summation in this final expression is the same as the summation in our explicit expression

for ψ, so we may use this equation to write

$$\begin{aligned}
&\psi(x_1,\ldots,x_{n-1})\\
&= \frac{n_k}{\sqrt{(n-1)!\prod_i n_i!}} \sum_{\sigma\in S_{n-1}} e_{m'_{\sigma(1)}}(x_1)\cdots e_{m'_{\sigma(n-1)}}(x_{n-1})\\
&= \frac{\sqrt{n}}{\sqrt{n!\prod_i n_i!}} \int_{-\infty}^{\infty} e_k(y)^* \sum_{\sigma\in S_n} e_{m_{\sigma(1)}}(x_1)\cdots e_{m_{\sigma(n-1)}}(x_{n-1})\, e_{m_{\sigma(n)}}(y)\, dy\\
&= \sqrt{n}\int_{-\infty}^{\infty} e_k(y)^*\varphi(x_1,\ldots,x_{n-1},y)\, dy.
\end{aligned}$$

Because we showed this equality for a basis vector $\varphi = |n_1, n_2, \ldots\rangle$, by linearity we have that for all $\varphi \in L^2(\mathbb{R}^n)_{sym}$,

$$(a_k\varphi)(x_1,\ldots,x_{n-1}) = \sqrt{n}\int_{-\infty}^{\infty} e_k(y)^*\varphi(x_1,\ldots,x_{n-1},y)\, dy.$$

Now, again using linearity, for any $f = \sum_{k=1}^{\infty}\alpha_k e_k \in \mathcal{H} = L^2(\mathbb{R})$, we have that

$$\begin{aligned}
(A(f)\varphi)(x_1,\ldots,x_{n-1}) &= \sum_k \alpha_k^*(a_k\varphi)(x_1,\ldots,x_{n-1})\\
&= \sqrt{n}\sum_k \alpha_k^*\int_{-\infty}^{\infty} e_k(y)^*\varphi(x_1,\ldots,x_{n-1},y)\, dy\\
&= \sqrt{n}\int_{-\infty}^{\infty} f(y)^*\varphi(x_1,\ldots,x_{n-1},y)\, dy.
\end{aligned}$$

This expression for $A(f)\varphi$ is clearly basis-independent. Although this expression is only for $\varphi \in L^2(\mathbb{R}^n)_{sym}$, by linearity we have that $A(f)\varphi$ is also basis-independent for general $\varphi \in \mathcal{B}$.

7.3. Basis-independence of $A^\dagger$

We take a similar approach for $A^\dagger(f)$. Again let $\varphi = |n_1, n_2, \ldots\rangle \in \mathcal{H}_{sym}^{\otimes n} = L^2(\mathbb{R}^n)_{sym}$, so that

$$\varphi(x_1,\ldots,x_n) = \frac{1}{\sqrt{n!\prod_i n_i!}} \sum_{\sigma\in S_n} e_{m_{\sigma(1)}}(x_1)\cdots e_{m_{\sigma(n)}}(x_n)$$

for some list of integers $m_1, m_2, \ldots, m_n$ for which n_i counts the number of i's listed. Let $\psi = a_k^\dagger\varphi \in L^2(\mathbb{R}^{n+1})_{sym}$. We know that

$$\psi = \sqrt{n_k+1}\,|n_1, n_2, \ldots, n_k+1, \ldots\rangle,$$

yielding

$$\begin{aligned}
&\psi(x_1,\ldots,x_{n+1})\\
&= \frac{\sqrt{n_k+1}}{\sqrt{(n+1)!(n_k+1)!\prod_{i\neq k} n_i!}} \sum_{\sigma\in S_{n+1}} e_{m'_{\sigma(1)}}(x_1)\cdots e_{m'_{\sigma(n+1)}}(x_{n+1}),
\end{aligned}$$

where $m'_1, m'_2, \ldots, m'_{n+1}$ is a list obtainted by adding a k to original list. Then it is not hard to see that

$$\begin{aligned}
&\sum_{\sigma\in S_{n+1}} e_{m'_{\sigma(1)}}(x_1)\cdots e_{m'_{\sigma(n+1)}}(x_{n+1}) \\
&= \sum_{j=1}^{n+1} e_k(x_j)\Bigg(\sum_{\sigma\in S_n} e_{m_{\sigma(1)}}(x_1)\cdots e_{m_{\sigma(j-1)}}(x_{j-1}) \\
&\qquad\qquad \cdot e_{m_{\sigma(j)}}(x_{j+1})\cdots e_{m_{\sigma(n)}}(x_{n+1})\Bigg).
\end{aligned}$$

Using this expression, we have

$$\begin{aligned}
&\psi(x_1,\ldots,x_{n+1}) \\
&= \frac{1}{\sqrt{(n+1)!\prod_i n_i!}} \sum_{\sigma\in S_{n+1}} e_{m'_{\sigma(1)}}(x_1)\cdots e_{m'_{\sigma(n+1)}}(x_{n+1}) \\
&= \frac{1}{\sqrt{(n+1)!\prod_i n_i!}} \sum_{j=1}^{n+1} e_k(x_j) \\
&\qquad \cdot \Bigg(\sum_{\sigma\in S_n} e_{m_{\sigma(1)}}(x_1)\cdots e_{m_{\sigma(j-1)}}(x_{j-1}) e_{m_{\sigma(j)}}(x_{j+1})\cdots e_{m_{\sigma(n)}}(x_{n+1})\Bigg) \\
&= \frac{1}{\sqrt{n+1}} \sum_{j=1}^{n+1} e_k(x_j)\, \varphi(x_1,\ldots,\widehat{x_j},\ldots,x_{n+1}),
\end{aligned}$$

where $\widehat{\cdot}$ indicates an omitted term. So extending by linearity to all $\varphi \in L^2(\mathbb{R}^n)_{sym}$, we have

$$(a_k^\dagger \varphi)(x_1,\ldots,x_{n+1}) = \frac{1}{\sqrt{n+1}} \sum_{j=1}^{n+1} e_k(x_j)\, \varphi(x_1,\ldots,\widehat{x_j},\ldots,x_{n+1}).$$

Finally, for $f = \sum_{k=1}^\infty \alpha_k e_k \in \mathcal{H} = L^2(\mathbb{R})$, we have that

$$\begin{aligned}
(A^\dagger(f)\varphi)(x_1,\ldots,x_{n+1}) &= \sum_k \alpha_k (a_k^\dagger \varphi)(x_1,\ldots,x_{n+1}) \\
&= \frac{1}{\sqrt{n+1}} \sum_k \alpha_k \sum_{j=1}^{n+1} e_k(x_j)\, \varphi(x_1,\ldots,\widehat{x_j},\ldots,x_{n+1}) \\
&= \frac{1}{\sqrt{n+1}} \sum_{j=1}^{n+1} f(x_j)\, \varphi(x_1,\ldots,\widehat{x_j},\ldots,x_{n+1}).
\end{aligned}$$

This expression for $A^\dagger(f)\varphi$ is basis-independent as well, making $A^\dagger$ itself basis-independent.

7.4. Commutation relations

To summarize, we have the following two basis-independent expressions for $A(f)$ and $A^\dagger(f)$:

$$(A(f)\varphi)(x_1,\dots,x_{n-1}) = \sqrt{n}\int_{-\infty}^{\infty} f(y)^*\varphi(x_1,\dots,x_{n-1},y)\,dy,$$
$$(A^\dagger(f)\varphi)(x_1,\dots,x_{n+1}) = \frac{1}{\sqrt{n+1}}\sum_{j=1}^{n+1} f(x_j)\,\varphi(x_1,\dots,\widehat{x_j},\dots,x_{n+1})$$

where $\varphi \in L^2(\mathbb{R}^n)_{sym}$. These expressions allow us to define $a(x)$ and $a^\dagger(x)$, continuous analogues of a_k and $a_k^\dagger$. Heuristically, for each $x \in \mathbb{R}$, we set

$$a(x) = A(\delta_x), \qquad a^\dagger(x) = A^\dagger(\delta_x),$$

meaning that

$$(a(x)\varphi)(x_1,\dots,x_{n-1}) = \sqrt{n}\varphi(x_1,\dots,x_{n-1},x),$$
$$(a^\dagger(x)\varphi)(x_1,\dots,x_{n+1}) = \frac{1}{\sqrt{n+1}}\sum_{j=1}^{n+1}\delta(x-x_j)\,\varphi(x_1,\dots,\widehat{x_j},\dots,x_{n+1}).$$

Under these definitions, we see that we may symbolically re-express $A(f)$ and $A^\dagger(f)$ as

$$A(f) = \int dx\, f(x)^* a(x), \qquad A^\dagger(f) = \int dx\, f(x)\, a^\dagger(x).$$

(Note that the definitions of a and $a^\dagger$ do not make sense directly, since L^2 functions cannot be evaluated pointwise, and the delta function is not in L^2. Therefore, to be precise, we must integrate these against test functions, yielding A and $A^\dagger$, respectively.)

With our original a_k and $a_k^\dagger$, it is not hard to verify the commutation relation

$$[a_k, a_l^\dagger] = \delta_{k,l}1.$$

To derive the continuous analogue, first let $f = \sum \alpha_k e_k$ and $g = \sum \beta_k e_k$, and recall that

$$\begin{aligned}[A(f), A^\dagger(g)] &= \left[\sum_{k=1}^{\infty}\alpha_k^* a_k, \sum_{l=1}^{\infty}\beta_l a_l^\dagger\right] \\ &= \sum_{k,l}\alpha_k^*\beta_l[a_k, a_l^\dagger] \\ &= \sum_{k,l}\alpha_k^*\beta_l\delta_{k,l}1 \\ &= (f,g)1.\end{aligned}$$

But using the symbolic expressions, we have

$$[A(f), A^\dagger(g)] = \int dx\, dy\, f(x)^* g(y)\, [a(x), a^\dagger(y)].$$

This gives us the commutation relation

$$[a(x), a^\dagger(y)] = \delta(x-y)1. \tag{7.1}$$

Similarly, we may derive that

$$[a(x), a(y)] = [a^\dagger(x), a^\dagger(y)] = 0$$

for all $x, y \in \mathbb{R}$. Jointly, these are all the commutation relations satisfied by the a and $a^\dagger$ operators.

This is where physics classes start, with the 'operators' $a(x)$ and $a^\dagger(x)$ defined at every point in space and satisfying the commutation relations. Instead, we have defined this concept rigorously, using operator-valued distributions.

7.5. Creation and annihilation on a general state space

Let us now extend the definitions of a and $a^\dagger$ to general $\mathcal{H} = L^2(X, d\lambda)$, where X is some measurable space and λ is a measure on X. First, the notion of a delta function on such a space is a distribution that maps a function f to its value at a specific point. That is,

$$\delta_x(f) = f(x) \qquad \forall x \in X.$$

Our definitions of A and $A^\dagger$ work for general $\mathcal{H}$, and in particular for $\mathcal{H} = L^2(X, d\lambda)$. We symbolically represent A and $A^\dagger$ as

$$A(f) = \int d\lambda(x) f(x)^* a(x)$$

$$A^\dagger(f) = \int d\lambda(x) f(x) a^\dagger(x).$$

This yields similar expressions for a and $a^\dagger$:

$$(a(x)\varphi)(x_1, \ldots, x_{n-1}) = \sqrt{n}\varphi(x_1, \ldots, x_{n-1}, x)$$

$$(a^\dagger(x)\varphi)(x_1, \ldots, x_{n+1}) = \frac{1}{\sqrt{n+1}} \sum_{j=1}^{n+1} \delta_{x_j}(x)\varphi(x_1, \ldots, \widehat{x_j}, \ldots, x_{n+1}).$$

LECTURE 8

Time evolution on Fock space

Date: 10/10/2018
Scribe: Andy Tsao

8.1. Defining $a^\dagger(x)a(x)$

Ordinarily, it does not make sense to talk about the product of two distributions. However, creation and annihilation operators can be multiplied in certain situations. We have already seen one example of that in the commutation relations for a and $a^\dagger$. Let us now see one more example. Given any smooth test function f, we will show that $\int dx f(x)a^\dagger(x)a(x)$ is a well-defined operator on $\mathcal{B}_0$. Indeed, take any $\varphi \in L^2(\mathbb{R}^n)$, and let $\psi = a(x)\varphi$ and $\xi = a^\dagger(x)\psi$. Then,

$$\psi(x_1, \ldots, x_{n-1}) = \sqrt{n}\varphi(x_1, \ldots, x_{n-1}, x),$$

and

$$\begin{aligned}\xi(x_1, \ldots, x_n) &= \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \delta(x - x_j)\psi(x_1, \ldots, \widehat{x_j}, \ldots, x_n) \\ &= \sum_{j=1}^{n} \delta(x - x_j)\varphi(x_1, \ldots, \widehat{x_j}, \ldots, x_n, x).\end{aligned}$$

Integrating over x gives us

$$\begin{aligned}\left(\int dx f(x)\xi\right)(x_1, \ldots, x_n) &= \sum_{j=1}^{n} \int dx f(x)\delta(x - x_j)\varphi(x_1, \ldots, \widehat{x_j}, \ldots, x_n, x) \\ &= \sum_{j=1}^{n} f(x_j)\varphi(x_1, \ldots, \widehat{x_j}, \ldots, x_n, x_j) \\ &= \left(\sum_{j=1}^{n} f(x_j)\right)\varphi(x_1, \ldots, x_n),\end{aligned}$$

where the last equality holds because φ is symmetric.

8.2. Free evolution on Fock space

We have shown previously that the free evolution Hamiltonian H acts as $H\varphi = -\frac{1}{2m}\Delta\varphi$ for $\varphi \in L^2(\mathbb{R}^n)_{sym}$. From this it follows by linearity that

the free evolution Hamiltonian on $\mathcal{B}_0$ is

$$H\left(\sum_{n=0}^{\infty}\varphi_n\right) = -\frac{1}{2m}\sum_{n=0}^{\infty}\Delta\varphi_n.$$

Physicists will say that

$$H = \int_{-\infty}^{\infty} dx a^\dagger(x)\left(-\frac{1}{2m}\frac{d^2}{dx^2}\right)a(x).$$

To see this, take $\varphi \in L^2(\mathbb{R}^n)_{sym}$ and let $\psi = a(x)\varphi$, $\xi = \frac{d^2}{dx^2}\psi$, and $\eta = a^\dagger(x)\xi$. Then,

$$\begin{aligned}
\psi(x_1,\ldots,x_{n-1}) &= \sqrt{n}\varphi(x_1,\ldots,x_{n-1},x),\\
\xi(x_1,\ldots,x_{n-1}) &= \sqrt{n}\partial_n^2\varphi(x_1,\ldots,x_{n-1},x),\\
\eta(x_1,\ldots,x_n) &= \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\delta(x-x_j)\xi(x_1,\ldots,\widehat{x_j},\ldots,x_n).
\end{aligned}$$

Integrating over x yields

$$\begin{aligned}
\left(\int dx\eta\right)(x_1,\ldots,x_n) &= \sum_{j=1}^{n}\int dx\delta(x-x_j)\partial_n^2\varphi(x_1,\ldots,\widehat{x_j},\ldots,x_n,x)\\
&= \sum_{j=1}^{n}\partial_n^2\varphi(x_1,\ldots,\widehat{x_j},\ldots,x_n,x_j)\\
&= \sum_{j=1}^{n}\partial_j^2\varphi(x_1,\ldots,x_n)\\
&= \Delta\varphi,
\end{aligned}$$

where the second-to-last equality follows from the symmetry of φ.

8.3. Introducing a potential

Consider n particles of mass m in a potential V. Then the Hamiltonian is adjusted to

$$H\varphi = -\frac{1}{2m}\Delta\varphi + \sum_{i=1}^{n}V(x_i)\varphi.$$

A similar argument to the above supports the representation

$$H = \int_{-\infty}^{\infty} dx a^\dagger(x)\left(-\frac{1}{2m}\frac{d^2}{dx^2} + V(x)\right)a(x).$$

Suppose further that any two particles repel each other with potential W. Then

$$H\varphi = -\frac{1}{2m}\Delta\varphi + \sum_{i}V(x_i)\varphi + \frac{1}{2}\sum_{i\neq j}W(x_i-x_j)\varphi$$

and has the representation

$$H = \int dx a^\dagger(x)\left(-\frac{1}{2m}\frac{d^2}{dx^2} + V(x)\right)a(x) + \int\int dxdy W(x-y)a^\dagger(x)a^\dagger(y)a(x)a(y).$$

Note that if $W(x-y) \sim \delta(x-y)$, the last term becomes close to

$$\int dx a^\dagger(x)a^\dagger(x)a(x)a(x),$$

and the whole thing can be written as an integral in x. However, we cannot set $W(x-y) = \delta(x-y)$ and still have the expression make sense in L^2.

8.4. Physical interpretation

Physicists will say that $a^\dagger(x)$ is an operator that creates a particle at x, while $a(x)$ is an operator that destroys a particle at x. To see this, suppose we take the state $\varphi = |0\rangle$. Let $\psi = a^\dagger(x)\varphi$ and $\xi = a^\dagger(y)\psi$. Then,

$$\begin{aligned}\psi(x_1) &= \frac{1}{\sqrt{0+1}}\delta(x-x_1)\varphi(\widehat{x_1}) = \delta(x-x_1),\\ \xi(x_1,x_2) &= \frac{1}{\sqrt{1+1}}(\delta(y-x_1)\psi(x_2) + \delta(y-x_2)\psi(x_1))\\ &= \frac{1}{\sqrt{2}}(\delta(y-x_1)\delta(x-x_2) + \delta(y-x_2)\delta(x-x_1)).\end{aligned}$$

We can see that ψ is the state with a particle located at x, and ξ is the symmetrized state with two particles located at x and y. Note that x and y need not be distinct.

8.5. Momentum space

Suppose that we now decide to represent all states as momentum wavefunctions. Then the conventional Hilbert space is $\mathcal{H} = L^2(\mathbb{R}, dp/2\pi)$. The operators a and $a^\dagger$ are almost the same as before, with

$$\begin{aligned}(a(p)\varphi)(p_1,\dots,p_{n-1}) &= \sqrt{n}\varphi(p_1,\dots,p_{n-1},p),\\ (a^\dagger(p)\varphi)(p_1,\dots,p_{n+1}) &= \frac{1}{\sqrt{n+1}}\sum_{j=1}^{n+1} 2\pi\delta(p-p_j)\varphi(p_1,\dots,\widehat{p_j},\dots,p_{n+1}).\end{aligned}$$

Note that an additional factor of 2π appears in the expression of $a^\dagger$ due to the scaled measure $dp/2\pi$.

Recall the Hamiltonian operator $H = -\frac{1}{2m}\Delta$ on position space. To emphasize that this is the Hamiltonian on position space, let us denote this by H_{x}. Let us now see what should be the analogous Hamiltonian operator on momentum space. If H_{p} denotes this operator, its defining property is that it must satisfy $H_{\mathrm{p}}\widehat{\varphi} = \widehat{H_{\mathrm{x}}\varphi}$ for all φ.

The n-fold tensor product of $\mathcal{H}$ is $L^2(\mathbb{R}^n, (2\pi)^{-n} dp_1 \cdots dp_n)$. For $\varphi \in \mathcal{H}^{\otimes n}$, the Hamiltonian operator on momentum space acts as

$$H_{\mathrm{p}}\varphi = \left(\frac{1}{2m} \sum_{j=1}^{n} p_j^2 \right) \varphi.$$

Since H_{p} acts on φ by scalar multiplication, its representation using a and $a^\dagger$ is simply

$$H_{\mathrm{p}} = \int \frac{dp}{2m} \frac{p^2}{2m} a^\dagger(p) a(p).$$

One can check that H_{p} and H_{x} satisfy the relationship $H_{\mathrm{p}}\widehat{\varphi} = \widehat{H_{\mathrm{x}}\varphi}$. In general we will not differentiate between H_{p} and H_{x} and denote them both by H (although they are operators on different spaces).

We will use the following notation throughout the rest of the lectures:

$$|p_1, p_2, \ldots, p_n\rangle := a^\dagger(p_1) a^\dagger(p_2) \cdots a^\dagger(p_n) \, |0\rangle \, .$$

Just like in position space, the above state is the state of n non-relativistic bosons in one-dimensional space with momenta exactly equal to $p_1, \ldots, p_n$.

Generally, we will be working in momentum space when we move to developing QFT. Consider a state $\psi = \sum_{n=0}^{\infty} \psi_n \in \mathcal{B}$, where $\mathcal{B}$ is the bosonic Fock space associated with the Hilbert space $L^2(\mathbb{R}, dp/2\pi)$. Then note that

$$\langle p_1, p_2, \ldots, p_n | \psi \rangle = \langle p_1, p_2, \ldots, p_n | \psi_n \rangle \, ,$$

since inner products of the form $\langle p_1, p_2, \ldots, p_n | \psi_m \rangle$ are zero when $n \neq m$ (states with different particle number are orthogonal by definition in the Fock space). This shows that $| \langle p_1, p_2, \ldots, p_n | \psi \rangle |^2$ is proportional to the joint probability density of the n momenta, conditional on the number of particles $= n$, if the state of the system is ψ. Recall that the system has n particles with probability proportional to $\|\psi_n\|^2$.

LECTURE 9

Special relativity

Date: 10/12/2018
Scribe: George Hulsey

9.1. Special relativity notation

We have now finished our preliminary discussion of non-relativistic quantum field theory. We will now move on to incorporate special relativity. As always, we will be adopting units where $\hbar = c = 1$. The following notations and conventions are standard in special relativity, and carry over to quantum field theory.

In special relativity, our universe is represented by the vector space $\mathbb{R}^{1,3}$, which is just $\mathbb{R}^4$ but with a different inner product. We will denote a vector $x \in \mathbb{R}^{1,3}$ without an arrow or bolding. We will write components as:

$$x = (x^0, x^1, x^2, x^3)$$

where x^0 is the time coordinate (note that we have used superscripts). Given $x \in \mathbb{R}^{1,3}$, we let $\mathbf{x} = (x^1, x^2, x^3)$ be the 3-tuple of spatial coordinates, which is a vector in $\mathbb{R}^3$. Then $x = (x^0, \mathbf{x})$. The distinction between x and $\mathbf{x}$ is very important.

We now define a symmetric matrix η as:

$$\eta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{9.1}$$

We will write $\eta_{\mu\nu}$ to denote the (μ, ν)th entry of η. Here, we use a subscript while we used a superscript to label the coordinates.

This matrix serves as a quadratic form to define an inner product on $\mathbb{R}^{1,3}$. Given $x, y \in \mathbb{R}^{1,3}$, the inner product is defined as:

$$(x, y) = x^0 y^0 - x^1 y^1 - x^2 y^2 - x^3 y^3 = \sum_{\mu,\nu} \eta_{\mu\nu} x^\mu y^\nu$$

This is the Minkowski inner product. But if we simply consider $\mathbf{x} \cdot \mathbf{y}$, we get the usual:

$$\mathbf{x} \cdot \mathbf{y} = x^1 y^1 + x^2 y^2 + x^3 y^3.$$

This is the regular dot product. So the Minkowski inner product can be written more succinctly as

$$(x, y) = x^0 y^0 - \mathbf{x} \cdot \mathbf{y}.$$

In special relativity, it is conventional to write $\mathbf{x}^2$ for $\mathbf{x} \cdot \mathbf{x}$ and

$$x^2 = (x, x) = (x^0)^2 - \mathbf{x} \cdot \mathbf{x}.$$

The Euclidean norm of a 3-vector $\mathbf{x}$ will be denoted by $|\mathbf{x}|$. It is vital to remember and be comfortable with all of the above notations for the rest of this lecture series.

9.2. Lorentz Transformations

A Lorentz transformation L on $\mathbb{R}^{1,3}$ is a linear map such that $(Lx, Ly) = (x, y)$ for all $x, y \in \mathbb{R}^{1,3}$ (using the Minkowski inner product). In the language of matrices, this means that:

$$L^T \eta \, L = \eta.$$

This is very similar to the orthogonality condition under the usual inner product. Like orthogonal matrices, Lorentz transformations form a Lie group $O(1, 3)$. In analogy with the regular orthogonal groups, we can consider the map $L \mapsto \det(L)$ where the determinant serves as a homomorphism:

$$\det : O(1, 3) \to \mathbb{Z}_2 = \{-1, 1\}.$$

In the case of the orthogonal groups, quotienting out by the action of this homomorphism is sufficient to reduce $O(4)$ to $SO(4)$. However, the group $O(1, 3)$ has another homomorphism $L \mapsto \text{sign}(L^0_0)$, where L^0_0 is the top left entry of the matrix.

We define the restricted Lorentz group $SO^{\uparrow}(1, 3)$ to be the set of all $L \in O(1, 3)$ such that $\det(L) = 1$ and $\text{sign}(L^0_0) = 1$. This is a subgroup of $O(1, 3)$.

9.3. What is special relativity?

Classical physics is invariant under isometries of $\mathbb{R}^3$, that is, translations and rotations. What does that mean? Suppose you have a computer program simulating physics. You input the state of the system at time 0, then the program gives you the state at time t. Suppose a trickster enters the state of the system at time 0 but in a different coordinate system. If your program is properly built, the result it returns for the trickster will (after changing back coordinates) be the exact same as the result you saw earlier. This a property of classical physics.

Special relativity claims that the laws of physics remain invariant under restricted Lorentz transformations and spacetime translations of $\mathbb{R}^{1,3}$, in the same sense as above. Now space and time become mixed into a single entity called spacetime, coordinates of which are changed by Lorentz transformations and spacetime translations. The same analogy as before holds,

except now the trickster can change the time coordinates under a Lorentz transformation with no effect on the final results.

Our goal in constructing a quantum field theory will be to find one invariant under these restricted Lorentz transformations and spacetime translations. We will do this, but we will not explicitly show that it is invariant. This is how the theories are designed, but this will not be proven to save time.

9.4. Proper time

The parametrization of a curve in Euclidean space by arc length has the property that it is invariant under Euclidean symmetries. That is, if we apply a rotation or translation to the coordinate system, the parametrization remains unchanged. In a similar way, a curve in $\mathbb{R}^{1,3}$ can sometimes be parametrized in a way that is invariant under any change of the coordinate system by an action of $SO^{\uparrow}(1,3)$ or a spacetime translation. The parametrization $x(\tau)$ of a curve by arc length requires that $\|dx/d\tau\| \equiv 1$, where $\|\cdot\|$ is Euclidean norm. In a similar vein, a Lorentz-invariant parametrization requires

$$\left(\frac{dx}{d\tau}, \frac{dx}{d\tau}\right) \equiv 1,$$

where $(\cdot,\cdot)$ is the Minkowski inner product. The conditions under which such a parametrization exists are a little more demanding than the conditions for parametrizability by arc length. A sufficient condition is that in some coordinate system, the curve has a time parametrization $x(t) = (t, \mathbf{x}(t))$ that satisfies

$$\left|\frac{d\mathbf{x}}{dt}\right| < 1 \quad \forall t.$$

In other words, in some coordinate system, the curve represents the trajectory of a particle whose speed is always strictly less than the speed of light.

If $x(\tau)$ is a parametrization of the trajectory of a particle as above, τ is called the 'proper time' of the trajectory. The proper time is (up to an additive constant) independent of the coordinate system under restricted Lorentz transformations and spacetime translations.

9.5. Four-momentum

Suppose that we are given the trajectory x of a particle, parametrized by proper time. If the mass of the particle is m, its 'four-momentum' at any point on the trajectory is the vector

$$p = m\frac{dx}{d\tau}.$$

By the nature of the proper time, p does not depend on the coordinate system. If you fix a coordinate system, then the four-momentum has the

formula

$$p = \left(\frac{m}{\sqrt{1 - \mathbf{v}(t)^2}}, \frac{m\mathbf{v}(t)}{\sqrt{1 - \mathbf{v}(t)^2}} \right),$$

where $\mathbf{v}(t) = d\mathbf{x}/dt$ is the coordinate-dependent velocity. We usually denote four-momentum as:

$$p = (p^0, p^1, p^2, p^3) = (p^0, \mathbf{p})$$

The four-momentum is also called the 'energy-momentum', as we can identify the first component of p^0 as the relativistic energy E. The remaining part, $\mathbf{p}$, is called the relativistic momentum. When the particle is at rest, $\mathbf{p} = 0$, and we get $E = m$, the famous Einstein relation (with $c = 1$). Further, note that in the non-relativistic limit $|\mathbf{v}(t)| \ll 1$, we can expand the energy p^0 as:

$$p^0 = \frac{m}{\sqrt{1 - \mathbf{v}(t)^2}} \approx m + \frac{1}{2} m\mathbf{v}(t)^2,$$

which is the sum of the rest energy and the classical kinetic energy. One thing we can observe from the definition of p is that:

$$p^2 = (p^0)^2 - \mathbf{p} \cdot \mathbf{p} = m^2,$$

which in turn implies that

$$p^0 = E = \sqrt{m^2 + \mathbf{p}^2} = \omega_{\mathbf{p}},$$

where the latter will become our conventional notation, with the subscript denoting the dependence on the relativistic momentum. Notice that the four-momentum p is always an element of the manifold

$$X_m = \{p \in \mathbb{R}^{1,3} : p^2 = m^2,\ p^0 \geq 0\}.$$

This manifold is called the 'mass shell' of mass m. It is a very important object in quantum field theory.

LECTURE 10

The mass shell

Date: 10/15/2018
Scribe: Alec Lau

10.1. Time evolution of a relativistic particle

Recall the mass shell $X_m = \{p \in \mathbb{R}^{1,3} : p^2 = m^2,\, p^0 \geq 0\}$ defined in the last lecture. In quantum field theory, we model the behavior of the four-momentum of a particle instead of the classical momentum. The Hilbert space is $\mathcal{H} = L^2(X_m, d\lambda_m)$, where λ_m is the unique measure (up to a multiplicative constant) that is invariant under the action of the restricted Lorentz group on X_m. We will define λ_m later in this lecture.

Suppose that this invariant measure exists. How does the state of the free particle evolve? The main postulate is the following:

Postulate. If the four-momentum state of a freely evolving particle at time 0 is $\psi \in L^2(X_m, d\lambda_m)$, then its state at time t is the function ψ_t, given by $\psi_t(p) = e^{-itp^0}\psi(p)$. (Recall that p^0 is the first coordinate of p.)

The reader may be wondering how time evolution makes sense in the above manner, when time itself is not a fixed notion in special relativity. Indeed, the above concept of time evolution does not make sense in the relativistic setting. The above postulate is simply the convenient way to think about what is going on. What is *really* going on is a bit more complicated. We will talk about it in the next lecture.

The above postulate is a direct generalization of free evolution in the non-relativistic setting in $\mathbb{R}^3$, since in that case $\psi_t = e^{-it\mathbf{p}^2/2m}\psi$, where $\mathbf{p}$ is the non-relativistic momentum and $\mathbf{p}^2/2m$ is the non-relativistic energy. The postulate simply replaces the non-relativistic energy by the relativistic energy p^0. Notice that the Hamiltonian for time evolution in the relativistic setting is therefore just $H\psi(p) = p^0\psi(p)$. Consequently,

$$\frac{\partial}{\partial t}\psi_t(p) = -ip^0\psi_t(p).$$

While this equation is actually correct, note that there is an obvious conceptual difficulty because the equation gives a special status to time, which is not acceptable in special relativity. We will resolve this difficulty in the next lecture.

10.2. The measure λ_m

We will now construct the invariant measure on our mass shell in analogy with the unique invariant measure on a sphere with respect to rotations. One way to construct the uniform measure on the sphere is to make a thin annulus around the sphere, take the Lebesgue measure on the annulus, normalize it, and take the width of the annulus to zero. The key to why this works is that a thin annulus is also rotationally invariant, giving a rotationally invariant measure in the limit.

One way to define Lorentz invariant annuli is to set

$$X_{m,\varepsilon} = \{p : m^2 < p^2 < (m+\varepsilon)^2\},$$

where the square is the Minkowski norm, hence making this annulus Lorentz invariant. Scaling Lebesgue measure on this annulus in a suitable way gives a nontrivial measure as $\varepsilon \to 0$. To integrate functions with respect to this measure, we bring our annuli down to $\mathbb{R}^3$. Any point $\mathbf{p} \in \mathbb{R}^3$ corresponds to a unique point $p \in X_m$ where $p = (\omega_{\mathbf{p}}, \mathbf{p})$, with $\omega_{\mathbf{p}} = \sqrt{m^2 + \mathbf{p}^2}$. The thickness of $X_{m,\varepsilon}$ at p is

$$\sqrt{(m+\varepsilon)^2 + \mathbf{p}^2} - \sqrt{m^2 + \mathbf{p}^2} \approx \frac{\varepsilon m}{\omega_{\mathbf{p}}}.$$

From this it is easy to derive that for an appropriate scaling of Lebesgue measure on $X_{m,\varepsilon}$ as $\varepsilon \to 0$, the scaling limit λ_m satisfies, for any integrable function f on X_m,

$$\int_{X_m} d\lambda_m(p) f(p) = \int_{\mathbb{R}^3} \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{p}}} f(\omega_{\mathbf{p}}, \mathbf{p}). \tag{10.1}$$

Note that constants do not matter because we are free to define our measure up to a constant multiple. The factor 2 in the denominator is conventional. The above integration formula gives a convenient way to integrate on the mass shell.

LECTURE 11

The postulates of quantum field theory

Date: 10/17/2018
Scribe: Henry Froland

11.1. Changing the postulates

We now arrive at a fundamental problem with our picture, which is, 'what does it mean to say momentum state at a given time?' This is because Lorentz transforms change the concept of time slices. That is, the notion of two events happening 'at the same time' need not remain the same under a change of coordinate system. As we have defined it, ψ evolving as $\psi_t(p) = e^{-itp^0}\psi(p)$ is just a convenient way to think of time evolution, but the true picture is more complicated.

To go from quantum mechanics to QFT, we need to fundamentally change the postulates to get a picture that is fully consistent with special relativity. Recall that we had five postulates of quantum mechanics that were introduced in the second lecture. The postulates need to be modified as follows.

- P1 For any quantum system, there exists a Hilbert space $\mathcal{H}$ such that the state of the system is described by vectors in $\mathcal{H}$. This looks the same as the original P1, but there is one crucial difference. In the new version, a state is not for a given time, but for all spacetime. In other words, a state gives a complete spacetime description of the system.
- P2 This postulate is unchanged.
- P3 This postulate is unchanged.
- P4 This postulate is unchanged.
- P5 This postulate is completely different. There is no concept of time evolution of a state in QFT. We need some extra preparation before stating this postulate, which we do below.

The Poincaré group $\mathcal{P}$ is the semi-direct product $\mathbb{R}^{1,3} \rtimes SO^{\uparrow}(1,3)$, where $\mathbb{R}^{1,3}$ is the group of spacetime translations (for $x \in \mathbb{R}^{1,3}, x(y) = x + y$). This means

$$\mathcal{P} = \{(a, A) : a \in \mathbb{R}^{1,3}, A \in SO^{\uparrow}(1,3)\} \tag{11.1}$$

with the group operation defined as $(a, A)(b, B) = (a + Ab, AB)$. This is the group of isometries of the Minkowski spacetime. The action of $(a, A) \in \mathcal{P}$ on $x \in \mathbb{R}^{1,3}$ is defined as $(a, A)(x) = a + Ax$.

A representation U of the Poincaré group in a Hilbert space $\mathcal{H}$ is a map U from $\mathcal{P}$ into the space of unitary operators on $\mathcal{H}$ such that

$$U((a, A)(b, B)) = U(a, A)U(b, B)$$

for all (a, A) and (b, B). The representation is called strongly continuous if for any $x \in \mathcal{H}$, the map $(a, A) \to U(a, A)x$ is continuous. The revised Postulate 5 replaces the notion of a time evolution group of unitary operators with a strongly continuous unitary representation of the Poincaré group:

P5 Any quantum system with Hilbert space $\mathcal{H}$ comes equipped with a strongly continuous unitary representation U of $\mathcal{P}$ in $\mathcal{H}$, which is related to the system in the following way: If an observer in a given coordinate system observes the system to be in state ψ, then an observer in a new coordinate system obtained by the action of (a, A) on $\mathbb{R}^{1,3}$ observes the state of the same system as $U(a, A)\psi$.

Actually, the correct version says 'projective unitary representation'; we will come to that later. The representation property provides consistency, meaning that if we change the reference frame two times in succession, the resulting state is the same as if we changed the frame once, obeying the composition law.

11.2. Implication for spacetime evolution

Let us now discuss what the new Postulate 5 means in terms of the time evolution of a system. Consider the transformation $(a, 1)$ acting on a point $x = (x^0, x^1, x^2, x^3)$, where $a = a(t) = (-t, 0, 0, 0)$ for some t. The result of this action gives us the point $(x^0 - t, x^1, x^2, x^3)$. This reduces the time coordinate by t, so the new reference frame obtained by the action of $(a, 1)$ is that of an observer who is t units ahead in time. This is an important point to understand. The new state $U(a, 1)\psi$ is what the system will look like t units of time later, if it looks like ψ now. In this sense, it is inaccurate to say that the state of a system is just an element of the Hilbert space. There is a full equivalence class that we obtain by the transformations, which is the abstract state of the system.

In some sense, we recover time evolution, but in actuality this is a different notion. Note that the family of unitary operators $(U(a(t), 1))_{t\in\mathbb{R}}$ is a strongly continuous unitary group acting on $\mathcal{H}$, so there exists a self-adjoint operator H such that $U(a(t), 1) = e^{-itH}$. If we fix a reference frame, then the system behaves just as if we were in the original quantum mechanical setting, with H being the Hamiltonian governing time evolution of states.

11.3. The representation for massive scalar bosons

Now we write down the representation for the particles that we have been considering until now (which are called massive scalar bosons). In our setting, $\mathcal{H} = L^2(X_m, d\lambda_m)$ (later on we will take the Fock space), and the

associated representation is

$$(U(a, A)\psi)(p) = e^{i(a,p)}\psi(A^{-1}p).$$

Taking $A = 1$ and $a = (-t, 0, 0, 0)$, we recover the time evolution group defined in the previous lecture, namely, that the state evolves as $\psi_t(p) = e^{-itp^0}\psi(p)$ in a fixed coordinate system.

LECTURE 12

The massive scalar free field

Date: 10/19/2018
Scribe: Manisha Patel

12.1. Creation and annihilation on the mass shell

Let $\mathcal{H} = L^2(X_m, d\lambda_m)$, and let $\mathcal{B}$ be the bosonic Fock space for this $\mathcal{H}$. Recall the operator-valued distributions A, $A^\dagger$ on the Fock space $\mathcal{B}$, and the formal representations

$$A(f) = \int d\lambda_m f^*(p) a(p),$$
$$A^\dagger(f) = \int d\lambda_m f(p) a^\dagger(p).$$

These are the usual, basis-independent definitions we have been working with. Recall that

$$(a(p)\phi)(p_1, \ldots, p_n) = \sqrt{n}\phi(p_1, \ldots, p_{n-1}, p),$$
$$(a^\dagger(p)\phi)(p_1, \ldots, p_{n+1}) = \frac{1}{\sqrt{n+1}} \sum_{j=1}^{n+1} \delta_{p_j}(p)\phi(p_1, \ldots, \widehat{p_j}, \ldots, p_{n+1}),$$

where δ_{p_j} is the Dirac delta on X_m at the point p_j. Note that we do not write $\delta(p - p_j)$ because X_m is not a vector space, and so $p - p_j$ may not belong to X_m. We define two related operator-valued distributions on $\mathcal{B}$:

$$a(\mathbf{p}) = \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} a(p), \quad a^\dagger(\mathbf{p}) = \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} a^\dagger(p).$$

Note that because we are on X_m, the last three coordinates $\mathbf{p}$ define the first, so that $p = (\omega_{\mathbf{p}}, \mathbf{p})$.

The following are the commutation relations for the operators defined above, easily derived using the commutation relations for $a(p)$ and $a^\dagger(p)$ that we already know from previous discussions:

$$[a(\mathbf{p}), a(\mathbf{p}')] = 0,$$
$$[a^\dagger(\mathbf{p}), a^\dagger(\mathbf{p}')] = 0,$$
$$[a(\mathbf{p}), a^\dagger(\mathbf{p}')] = (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{p}')1.$$

For example, to prove the last identity, notice that by (7.1) and (10.1),

$$\begin{aligned}
&\iint \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{d^3\mathbf{p}'}{(2\pi)^3} f(\mathbf{p})^* g(\mathbf{p}')[a(\mathbf{p}), a^\dagger(\mathbf{p}')] \\
&= \iint \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{d^3\mathbf{p}'}{(2\pi)^3}\frac{1}{\sqrt{4\omega_{\mathbf{p}}\omega_{\mathbf{p}'}}} f(\mathbf{p})^* g(\mathbf{p}')[a(p), a^\dagger(p')] \\
&= \iint d\lambda_m(p) d\lambda_m(p')\sqrt{4\omega_{\mathbf{p}}\omega_{\mathbf{p}'}} f(\mathbf{p})^* g(\mathbf{p}')[a(p), a^\dagger(p')] \\
&= \left(\int d\lambda_m(p) 2\omega_{\mathbf{p}} f(\mathbf{p})^* g(\mathbf{p})\right) 1 \\
&= \left(\int \frac{d^3\mathbf{p}}{(2\pi)^3} f(\mathbf{p})^* g(\mathbf{p})\right) 1,
\end{aligned}$$

where 1 denotes the identity operator on $L^2(X_m, d\lambda_m)$.

12.2. The massive scalar free field

The massive scalar free field φ is an operator-valued distribution acting on $\mathscr{S}(\mathbb{R}^{1,3})$, the space of Schwartz functions on $\mathbb{R}^{1,3}$. The action of φ on a Schwartz function f is defined as

$$\varphi(f) = A(\widehat{f^*}\big|_{X_m}) + A^\dagger(\widehat{f}\big|_{X_m})$$

where the hat notation denotes the Fourier transform of f. On Minkowski space, the Fourier transform is defined as

$$\widehat{f}(p) = \int d^4x\, e^{i(x,p)} f(x).$$

Note that there is no minus sign in the exponent because we have the Minkowski inner product, so the minus sign is contained in the space coordinates.

The free field φ is the first real quantum field that we are seeing in this course. Quantum fields are operator-valued distributions. A field is an abstract 'thing' that doesn't exist as an object, but has real, observable effects. For example consider the classical magnetic field. For this field, we assign to each point in space a vector that denotes the direction and strength of the field at that point. Similarly we can also consider fields that put a scalar at each point. These are called classical scalar fields. Quantum mechanics replaces observables with operators, so this is how we arrive at an operator at each point in our spacetime. These fields act on particles by Hamiltonians defined using the fields.

12.3. Multiparticle states

We previously defined the state for a system with n particles with four-momenta exactly equal to $p_1, \ldots, p_n$ as

$$|p_1, \ldots, p_n\rangle = a^\dagger(p_1)\cdots a^\dagger(p_n)|0\rangle.$$

We now define

$$|\mathbf{p}_1, \ldots, \mathbf{p}_n\rangle = a^\dagger(\mathbf{p}_1) \cdots a^\dagger(\mathbf{p}_n)|0\rangle,$$

where $p = (p^0, \mathbf{p})$. This is the same state as $|p_1, \ldots, p_n\rangle$, up to a constant multiple. Conversely, if $\mathbf{p} \in \mathbb{R}^3$, we will adopt the convention that $p = (\omega_{\mathbf{p}}, \mathbf{p}) \in X_m$.

Consider a normalized state $\psi = \sum_{n=0}^{\infty} \psi_n \in \mathcal{B}$. We noted that conditional on the event that the number of particles $= n$ (which happens with probability $\|\psi_n\|^2$), the joint probability density of the n four-momenta, with respect to the invariant measure on X_m^n, is proportional to $|\langle p_1, \ldots, p_n|\psi\rangle|^2$. Suppose we want to find the joint density of the relativistic momenta (not the four-momenta) with respect to the Lebesgue measure on $\mathbb{R}^3$. (Recall that the relativistic momentum of a particle, often just called the momentum, is simply the 3-vector formed by the last three coordinates of the four-momentum.) We claim that this probability density is proportional to $|\langle \mathbf{p_1}, \ldots, \mathbf{p_n}|\psi\rangle|^2$.

To see why, consider a one-particle system. Take some region $A \subseteq \mathbb{R}^3$ and suppose that the system is in a normalized state $|\psi\rangle$. Let B be the image of A in X_m under the map $\mathbf{p} \mapsto p = (\omega_{\mathbf{p}}, \mathbf{p})$. Then

$$\begin{aligned}
\text{Prob}(\text{momentum} \in A) &= \text{Prob}(\text{four-momentum} \in B) \\
&= \int_B d\lambda_m(p) |\langle p|\psi\rangle|^2 \\
&= \int_A \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{|\langle p|\psi\rangle|^2}{2\omega_{\mathbf{p}}} \\
&= \int_A \frac{d^3\mathbf{p}}{(2\pi)^3} |\langle \mathbf{p}|\psi\rangle|^2
\end{aligned}$$

where in line 3 to 4, we used the fact that

$$|\mathbf{p}\rangle = a^\dagger(\mathbf{p})|0\rangle = \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} a^\dagger(p)|0\rangle = \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} |p\rangle.$$

12.4. Hamiltonian on Fock space

Let the Hamiltonian for free evolution on $\mathcal{H} = L^2(X_m, d\lambda_m)$ be denoted by H. We know that $H\psi(p) = p^0\psi(p)$. We will denote this H by H_0 henceforth, since we will work with more general Hamiltonians which will be denoted by H. Then the action of H_0 on $\mathcal{H}_{sym}^{\otimes n}$ is given by

$$H_0\psi(p_1, \cdots, p_n) = \left(\sum_{j=1}^{n} p_j^0\right)\psi(p_1, \cdots, p_n).$$

This extends to the dense subspace $\mathcal{B}_0$ of $\mathcal{B}$ by linearity.

Recall that on $L^2(\mathbb{R})$, we showed that the Hamiltonian H can be expressed formally as $\int_{-\infty}^{\infty} dp\ (p^2/2m) a^\dagger(p)a(p)$. A similar computation shows

that the above H_0 can be formally expressed as

$$\begin{aligned} H_0 &= \int_{X_m} d\lambda_m(p)\; p^0 a^\dagger(p) a(p) \\ &= \int_{\mathbb{R}^3} \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{p}}} \omega_{\mathbf{p}} \sqrt{2\omega_{\mathbf{p}}}\; a^\dagger(\mathbf{p}) \sqrt{2\omega_{\mathbf{p}}}\; a(\mathbf{p}) \\ &= \int_{\mathbb{R}^3} \frac{d^3\mathbf{p}}{(2\pi)^3} \omega_{\mathbf{p}} a^\dagger(\mathbf{p}) a(\mathbf{p}). \end{aligned}$$

12.5. Pointwise representation of the free field

Let φ be the scalar free field of mass m. Recall that for $f \in \mathscr{S}(\mathbb{R}^{1,3})$, the Fourier transform is defined as $\widehat{f}(p) = \int d^4x\, e^{i(x,p)} f(x)$. Thus,

$$\begin{aligned} \varphi(f) &= A\left(\widehat{f^*}|_{X_m}\right) + A^\dagger\left(\widehat{f}|_{X_m}\right) \\ &= \int d\lambda_m(p) \left(\widehat{f^*}(p)^* a(p) + \widehat{f}(p) a^\dagger(p)\right) \\ &= \int_{\mathbb{R}^3} \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{p}}} \left(\widehat{f^*}(p)^* a(p) + \widehat{f}(p) a^\dagger(p)\right) \\ &= \int_{\mathbb{R}^3} \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \left(\widehat{f^*}(p)^* a(\mathbf{p}) + \widehat{f}(p) a^\dagger(\mathbf{p})\right). \end{aligned}$$

After plugging in the integrals for $\widehat{f}(p)$ and $\widehat{f^*}(p)$, and exchanging integrals, we get

$$\varphi(f) = \int_{\mathbb{R}^{1,3}} d^4x f(x) \left[\int_{\mathbb{R}^3} \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \left(e^{-i(x,p)} a(\mathbf{p}) + e^{i(x,p)} a^\dagger(\mathbf{p})\right)\right].$$

The quantity inside $[\cdots]$ is called $\varphi(x)$. This is the pointwise representation of φ. We will use it extensively.

12.6. Normal ordering

Whenever we have a monomial of a's and $a^\dagger$'s, the normal ordered (or Wick ordered) version is the same monomial but with all $a^\dagger$'s brought to the left. The normal ordered monomial is denoted by putting colons in front and at the end. For example,

$$:a(\mathbf{p}) a^\dagger(\mathbf{p}): = a^\dagger(\mathbf{p}) a(\mathbf{p})$$

and

$$:a(\mathbf{p}_1) a^\dagger(\mathbf{p}_2) a^\dagger(\mathbf{p}_3) a(\mathbf{p}_4): = a^\dagger(\mathbf{p}_2) a^\dagger(\mathbf{p}_3) a(\mathbf{p}_1) a(\mathbf{p}_4).$$

Normal ordering will play an important role in the construction of interaction Hamiltonians later.

12.7. Expressing the Hamiltonian using the free field

Write $x = (t, \mathbf{x})$, where $\mathbf{x} = (x^1, x^2, x^3)$. Returning to our expression for $\varphi(x)$ derived previously, where

$$\varphi(x) = \int_{\mathbb{R}^3} \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \left(e^{-i(x,p)} a(\mathbf{p}) + e^{i(x,p)} a^\dagger(\mathbf{p}) \right),$$

we can formally differentiate $\varphi(x)$ with respect to t, x^1, x^2, and x^3, denoted by $\partial_t\varphi$, $\partial_1\varphi$, $\partial_2\varphi$, and $\partial_3\varphi$. It can be shown by a simple but slightly tedious calculation that for any t,

$$H_0 = \frac{1}{2} \int_{\mathbb{R}^3} d^3\mathbf{x} : \left((\partial_t\varphi(x))^2 + \sum_{\nu=1}^{3} (\partial_\nu\varphi(x))^2 + m^2\varphi(x)^2 \right) :.$$

LECTURE 13

Introduction to φ^4 theory

Date: 10/22/2018
Scribe: Youngtak Sohn

13.1. Evolution of the massive scalar free field

In the last lecture, we have defined the free field $\varphi(x)$ for $x \in \mathbb{R}^{1,3}$. Then $\varphi(x)$ is like an operator on $\mathcal{B}$ for any x. (More precisely it is an operator when one averages over a test function.) The following proposition describes how the massive free field evolves.

PROPOSITION 13.1. *For any $\mathbf{x} \in \mathbb{R}^3$ and any $t \in \mathbb{R}$,*

$$\varphi(t, \mathbf{x}) = e^{itH_0}\varphi(0, \mathbf{x})e^{-itH_0},$$

where H_0 is the free evolution Hamiltonian.

Recall the Heisenberg picture where a Hamiltonian H makes an operator B evolve as $B_t = e^{itH}Be^{-itH}$. The above proposition says that for any $\mathbf{x}$, $\varphi(t, \mathbf{x})$ evolves according to H_0. To prove the proposition, we start with the following lemma.

LEMMA 13.1. *If U is a unitary operator on $\mathcal{H}$, extended to $\mathcal{B}$, then for all $f \in \mathcal{H}$,*

$$UA(f)U^{-1} = A(Uf)$$
$$UA^\dagger(f)U^{-1} = A^\dagger(Uf).$$

PROOF. For $\mathcal{H} = L^2(\mathbb{R})$, write down $(UA(f)U^{-1})(g)$ explicitly using the formula for $A(f)$ when $g = g_1 \otimes \cdots \otimes g_n$, and check that it is the same as $A(Uf)(g)$. Similarly, check for $A^\dagger$. Extend to general $\mathcal{H}$ by isometries or a direct rewriting of the proof. □

PROOF OF PROPOSITION 13.1. The proof given below is formal, but can be made completely rigorous. Notice that

$$a(\mathbf{p}) = \frac{1}{\sqrt{2\omega_{\mathbf{p}}}}a(p) = \frac{1}{\sqrt{2\omega_{\mathbf{p}}}}A(\delta_p).$$

This implies

$$\begin{aligned} e^{itH_0}a(\mathbf{p})e^{-itH_0} &= \frac{1}{\sqrt{2\omega_{\mathbf{p}}}}e^{itH_0}A(\delta_p)e^{-itH_0} \\ &= \frac{1}{\sqrt{2\omega_{\mathbf{p}}}}A(e^{itH_0}\delta_p) \\ &= \frac{1}{\sqrt{2\omega_{\mathbf{p}}}}A(e^{itp^0}\delta_p) = \frac{e^{-itp^0}}{\sqrt{2\omega_{\mathbf{p}}}}A(\delta_p) = e^{-itp^0}a(\mathbf{p}). \end{aligned} \tag{13.1}$$

Similarly,

$$e^{itH_0}a^\dagger(\mathbf{p})e^{-itH_0} = e^{itp^0}a^\dagger(\mathbf{p}). \tag{13.2}$$

Then recall that

$$\varphi(t,\mathbf{x}) = \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\mathbf{p}}}}\left(e^{-i(x,p)}a(\mathbf{p}) + e^{i(x,p)}a^\dagger(\mathbf{p})\right),$$

where $(x,p) = tp^0 - \mathbf{x}\cdot\mathbf{p}$. Together with (13.1) and (13.2), this completes the proof. □

13.2. φ^4 theory

Recall that

$$H_0 = \frac{1}{2}\int d^3\mathbf{x} :\left((\partial_t\varphi(x))^2 + \sum_{\nu=1}^{3}(\partial_\nu\varphi(x))^2 + m^2\varphi(x)^2\right):,$$

for any given t, and in particular, $t=0$. Let H be the Hamiltonian H_0+gH_I, where g is the coupling parameter, which is typically small, and H_I is the interaction Hamiltonian, given by

$$H_I = \frac{1}{4!}\int d^3\mathbf{x} :\varphi(0,\mathbf{x})^4:.$$

So H is given by

$$\begin{aligned} H &= H_0 + gH_I \\ &= \int d^3\mathbf{x} :\left(\frac{1}{2}(\partial_t\varphi(0,\mathbf{x}))^2 + \frac{1}{2}\sum_{\nu=1}^{3}(\partial_\nu\varphi(0,\mathbf{x}))^2\right. \\ &\qquad \left. + \frac{m^2}{2}\varphi(0,\mathbf{x})^2 + \frac{g}{4!}\varphi(0,\mathbf{x})^4\right): . \end{aligned}$$

The motivation to consider the particular perturbation Hamiltonian is the following. Recall that in the one-dimensional non-relativistic setting, the Hamiltonian to describe n particles with mass m in a potential V and repelling each other with potential W is given by the following.

$$H\psi = -\frac{1}{2m}\frac{d^2}{dx^2}\psi + \sum_i V(x_i)\psi + \frac{1}{2}\sum_{i\neq j}W(x_i - x_j)\psi.$$

It also has the formal representation given by

$$H = \int dx\, a^\dagger(x) \left(-\frac{1}{2m}\frac{d^2}{dx^2} + V(x) \right) a(x) + \int\int dxdy\, W(x-y) a^\dagger(x) a^\dagger(y) a(x) a(y).$$

When $W(x-y)$ approaches $\delta(x-y)$, the last term becomes close to

$$\int dx\, a^\dagger(x) a^\dagger(x) a(x) a(x),$$

and the integrand is given by the normal ordering of four operator-valued distributions.

Up to this point in this class, everything was rigorous. However, in φ^4 theory, there is a problem of finding a suitable Hilbert space to make it a rigorous theory, despite the fact that one can carry out the calculation to make predictions. Say $\psi \in \mathcal{H}_{sym}^{\otimes n}$. Then we can write down $H_I\psi$ formally, but it does not need to be in $\mathcal{B}$. Thus, it leads to the following open question.

OPEN PROBLEM 13.1. Make sense of the Hamiltonian for φ^4 theory as a true Hamiltonian on a Hilbert space.

We want to know how states evolve according to H. To understand this, we will first have to learn about scattering theory, beginning in the next lecture.

Incidentally, φ^4 theory does not correspond to any real phenomena. The main purpose of analyzing φ^4 theory is that it possesses many of the main complexities of the more realistic models, and gives us an opportunity to develop various computational tools, such as Wick's theorem, Feynman diagrams, and renormalization.

In the next two lectures, we will move away from quantum field theory and return to non-relativistic quantum mechanics, to discuss the phenomenon of scattering. We will return to QFT after developing some necessary tools for understanding scattering in QFT.

LECTURE 14

Scattering

14.1. Classical scattering

Let V be a potential which is strong near $\mathbf{0} \in \mathbb{R}^3$, but very weak as you go away from $\mathbf{0}$. Let us first try to make sense of the following question in the setting of classical Newtonian mechanics: Suppose a particle moves towards the origin with velocity $\mathbf{v}$ under the influence of the potential V. What is the outgoing velocity?

The trajectory of a free particle is always of the form $(\mathbf{x}+t\mathbf{v})_{t\in\mathbb{R}}$, where $\mathbf{x}, \mathbf{v} \in \mathbb{R}^3$ and $\mathbf{x}$ denotes the position at time 0. Denote such a trajectory by $(\mathbf{x}, \mathbf{v})$. Given some trajectory $(\mathbf{x}, \mathbf{v})$, and some $t < 0$, consider a particle that is on this trajectory at time t. Let $\mathbf{x}'$ be its location at time 0 if it is moving under the influence of the potential V from time t onwards, and let $\mathbf{v}'$ be its velocity at time 0. Let $(\mathbf{x}, \mathbf{v})_t := (\mathbf{x}', \mathbf{v}')$ and define

$$\Omega_+(\mathbf{x}, \mathbf{v}) = \lim_{t\to-\infty} (\mathbf{x}, \mathbf{v})_t,$$

assuming that the limit exists. Then $\Omega_+(\mathbf{x}, \mathbf{v})$ can be interpreted as the (location, velocity) at time 0 of a particle coming in 'along the trajectory ($\mathbf{x}$,$\mathbf{v}$)' from the far past and moving under the influence of V.

Next, take $t > 0$ and look at a particle on the trajectory ($\mathbf{x}$,$\mathbf{v}$) at time t. Find $(\mathbf{x}', \mathbf{v}')$ such that if a particle were at $(\mathbf{x}', \mathbf{v}')$ at time 0, and the potential is turned on, it would be at $\mathbf{x}+t\mathbf{v}$ at time t. Here we assume that such a pair $(\mathbf{x}', \mathbf{v}')$ exists. Let $(\mathbf{x}, \mathbf{v})_t := (\mathbf{x}', \mathbf{v}')$ and define

$$\Omega_-(\mathbf{x}, \mathbf{v}) = \lim_{t\to\infty} (\mathbf{x}, \mathbf{v})_t,$$

again assuming that the limit exists. Then $\Omega_-(\mathbf{x}, \mathbf{v})$ is the (location,velocity) of a particle at time 0, which when moving in the potential, assumes the trajectory $(\mathbf{x}, \mathbf{v})$ in the far future.

Finally, the scattering operator is defined as

$$S := \Omega_-^{-1}\Omega_+,$$

if it makes sense. To understand what it means, let $(\mathbf{y}, \mathbf{u}) = S(\mathbf{x}, \mathbf{v})$. Then

$$\Omega_-(\mathbf{y}, \mathbf{u}) = \Omega_+(\mathbf{x}, \mathbf{v}).$$

The right hand side gives the (location, velocity) at time 0 if $(\mathbf{x}, \mathbf{v})$ is the trajectory in the far past. The left hand side gives the (location, velocity) at time 0 if $(\mathbf{y}, \mathbf{u})$ is the trajectory in the far future. This implies $(\mathbf{y}, \mathbf{u})$ is the trajectory in the far future if $(\mathbf{x}, \mathbf{v})$ is the trajectory in the far past.

14.2. Scattering in non-relativistic QM

We now know what scattering means in the classical case, although it is a bit complicated even in that context. Now, consider the setting of non-relativistic quantum mechanics in the context of a single particle in three dimensional space. Let

$$H = H_0 + gH_I$$

where H_0 is the free Hamiltonian, H_I some interaction Hamiltonian (e.g. $H_I = V$ for some potential V), and g is a coupling parameter. We want to understand evolution under H using the scattering approach. Let $U(t) = e^{-itH}$ and $U_0(t) = e^{-itH_0}$. If $|\psi\rangle$ is a state of the system at time 0, then its free evolution is the collection of states

$$(U_0(t)\,|\psi\rangle)_{t\in\mathbb{R}},$$

which is the analog of a straight line in the classical case. Let us identify $|\psi\rangle$ with this trajectory and call it 'the trajectory $|\psi\rangle$'. Now suppose that the state of the particle is evolving according to U instead of U_0. Also, suppose that it is 'in the trajectory $|\psi\rangle$' at some time $t < 0$. That is, its state at time t is $U_0(t)\,|\psi\rangle$. Then its state at time 0 is

$$U(-t)U_0(t)\,|\psi\rangle\,.$$

Define

$$\Omega_+\,|\psi\rangle \equiv \lim_{t\to-\infty} U(-t)U_0(t)\,|\psi\rangle$$

which is the 'state of a particle at time 0 if it is on the trajectory $|\psi\rangle$ in the far past'. Similarly, Ω_- is the state at time 0 of a particle that is on the trajectory $|\psi\rangle$ in the far future:

$$\Omega_-\,|\psi\rangle \equiv \lim_{t\to\infty} U(-t)U_0(t)\,|\psi\rangle\,.$$

As in the classical case, define the scattering operator

$$S = \Omega_-^{-1}\Omega_+.$$

That is, $S\,|\psi\rangle$ is the 'trajectory of a particle in the far future if it is on the trajectory $|\psi\rangle$ in the far past'.

If $|\varphi\rangle = S\,|\psi\rangle$, then $\Omega_-\,|\varphi\rangle = \Omega_+\,|\psi\rangle$. This means that 'if the particle looked like it was evolving as $U_0(t)\,|\psi\rangle$ for $t \ll 0$, then it will evolve as $U_0(t)\,|\varphi\rangle$ for $t \gg 0$'. More compactly,

$$S = \lim_{t_2\to\infty,\, t_1\to-\infty} U_0(-t_2)U(t_2 - t_1)U_0(t_1)$$

But there are two main problem in this set-up:

- Limits may not exist in the definitions of Ω_+ and Ω_-.
- We need $Range(\Omega_+) \subseteq Range(\Omega_-)$ to define $S = \Omega_-^{-1}\Omega_+$.

The condition $Range(\Omega_+) = Range(\Omega_-)$ is called 'asymptotic completeness'. If this is not valid, a particle can get 'trapped by the potential', and will not look like it is in a free state at late times. It is a complicated technical condition, and we will not bother to verify it because our main goal is

to move on to quantum field theory and essentially nothing can be verified rigorously in that setting.

Here is an example where the limits exist:

THEOREM 14.1. *Let $\mathcal{H} = L^2(\mathbb{R}^3)$, with the one-particle free Hamiltonian $H_0 = -\frac{1}{2m}\Delta$ and $H_I = V \in L^2(\mathbb{R}^3)$. Then, $H = H_0 + V$ is essentially self-adjoint, and Ω_+ and Ω_- exist as operators in $L^2(\mathbb{R}^3)$.*

PROOF SKETCH FOR Ω_+. Note that $U(-t)U_0(t)$ is a unitary operator. Thus,

$$\forall \psi, \varphi,\ \|U(-t)U_0(t)\varphi - U(-t)U_0(t)\psi\|_{L^2} = \|\varphi - \psi\|_{L^2},$$

is independent of t. So, it suffices to prove that $\lim_{t\to-\infty} U(-t)U_0(t)\varphi$ exists in L^2 for a spanning set of φ's. For example, it is enough to show this for

$$\varphi(\mathbf{x}) = e^{-(\mathbf{x}-\mathbf{a})^2/2\alpha},\ \mathbf{a} \in \mathbb{R}^3, \alpha > 0.$$

Explicitly solving the Schrödinger equation for such φ, it is not hard to show that

$$\|U_0(t)\varphi\|_{L^\infty} \leq \left(1 + \frac{t^2}{m^2\alpha^2}\right)^{-\frac{3}{4}}.$$

Let

$$\begin{aligned}\psi_t &= \frac{d}{dt}(U(-t)U_0(t)\varphi) = \frac{d}{dt}(e^{itH}e^{-itH_0}\varphi)\\ &= e^{itH}iHe^{-itH_0}\varphi - e^{itH}iH_0e^{-itH_0}\varphi\\ &= ie^{itH}(H - H_0)e^{-itH_0}\varphi = ie^{itH}Ve^{-itH_0}\varphi.\end{aligned}$$

Hence,

$$\begin{aligned}\|\psi_t\|_{L^2} &= \|ie^{itH}Ve^{-itH_0}\varphi\|_{L^2} = \|Ve^{-itH_0}\varphi\|_{L^2}\\ &\leq \|V\|_{L^2}\|e^{-itH_0}\varphi\|_{L^\infty} \leq \left(1 + \frac{t^2}{m^2\alpha^2}\right)^{-\frac{3}{4}}\|V\|_{L^2}.\end{aligned}$$

Now,

$$U(-t)U_0(t)\varphi = \varphi - \int_t^0 \psi_s ds.$$

So we have to show that

$$\lim_{t\to-\infty}\int_t^0 \psi_s ds$$

exists in L^2. It suffices to show that $\int_t^0 \|\psi_s\|_{L^2} ds < \infty$. This holds, because

$$\|\psi_s\|_{L^2} \leq \|V\|_{L^2}\left(1 + \frac{s^2}{m^2\alpha^2}\right)^{-\frac{3}{4}} \sim_{s\to\infty} \mathcal{O}(|s|^{-\frac{3}{2}}).$$

□

14.3. Dyson series expansion

The Dyson series expansion is a formal power series expansion for the scattering operator. Let H_0 and H_I be self-adjoint operators on a Hilbert space, and let

$$H = H_0 + gH_I.$$

Define the scattering operator

$$S = \lim_{t_0 \to -\infty,\, t \to +\infty} U_0(-t)U(t - t_0)U_0(t_0),$$

where $U(t) = e^{-itH}$ and $U_0(t) = e^{-itH_0}$. Fix $t_0 < 0$ and let

$$G(t) = U_0(-t)U(t - t_0)U_0(t_0).$$

Then a simple calculation gives

$$\frac{d}{dt}G(t) = -igH_I(t)G(t)$$

where $H_I(t) = U_0(-t)H_IU_0(t)$, $H_I(t)$ is the free evolution of H_I in the Heisenberg picture. Thus,

$$\begin{aligned} G(t) &= G(t_0) + \int_{t_0}^{t} dsG'(s) \\ &= G(t_0) - ig\int_{t_0}^{t} dsH_I(s)G(s) \\ &= 1 - ig\int_{t_0}^{t} dsH_I(s)G(s). \end{aligned}$$

Iterating, we get the formal power series:

$$G(t) = 1 + \sum_{n=1}^{\infty}(-ig)^n \int_{t_0}^{t}\int_{t_0}^{\theta_1}\cdots\int_{t_0}^{\theta_{n-1}} H_I(\theta_1)H_I(\theta_2)\cdots H_I(\theta_n)d\theta_n\cdots d\theta_1$$

Now send $t \to \infty$, and then $t_0 \to -\infty$, to get

$$\begin{aligned} S &= 1 + \sum_{n=1}^{\infty}(-ig)^n \int_{-\infty}^{\infty}\int_{-\infty}^{\theta_1}\cdots\int_{-\infty}^{\theta_{n-1}} H_I(\theta_1)H_I(\theta_2)\cdots H_I(\theta_n)d\theta_n\cdots d\theta_1 \\ &= 1 + \sum_{n=1}^{\infty}\frac{(-ig)^n}{n!}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \mathcal{T}H_I(\theta_1)H_I(\theta_2)\cdots H_I(\theta_n)d\theta_n\cdots d\theta_1, \end{aligned}$$

where the time ordering operator $\mathcal{T}$ is defined as:

$$\mathcal{T}H_I(\theta_1)H_I(\theta_2)\cdots H_I(\theta_n) = H_I(\theta_{\sigma(1)})H_I(\theta_{\sigma(2)})\cdots H_I(\theta_{\sigma(n)}),$$

where σ is a permutation such that $\theta_{\sigma(1)} \geq \theta_{\sigma(2)} \geq \cdots \geq \theta_{\sigma(n)}$. This is known as the Dyson series expansion of S. Note that rigorously, we can hope to prove a valid asymptotic series expansion, but actual convergence is unlikely.

LECTURE 15

The Born approximation

Date: 10/26/2018
Scribe: Julien Boussard

15.1. Derivation of the Born approximation

Suppose that we have a Hamiltonian $H = H_0 + gH_I$. Then we derived the Dyson series expansion for the scattering operator

$$S = 1 + \sum_{n=1}^{\infty} \frac{(-ig)^n}{n!} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dt_1 \cdots dt_n \mathcal{T} H_I(t_1) \cdots H_I(t_n),$$

where $H_I(t) = e^{itH_0} H_I e^{-itH_0}$ and

$$\mathcal{T} H_I(t_1) \cdots H_I(t_n) = H_I(t_{\sigma(1)}) \cdots H_I(t_{\sigma(n)})$$

for $\sigma \in S_n$ such that $t_{\sigma(1)} \geq \cdots \geq t_{\sigma(n)}$.

Now let $H_0 = -\frac{1}{2m}\Delta$ and $H_I = V \in L^2(\mathbb{R}^3)$. For this problem, we will now work out the first order approximation of S:

$$S = 1 + (-ig) \int_{-\infty}^{\infty} dt e^{itH_0} V e^{-itH_0} + \mathcal{O}(g^2).$$

Let us try to compute $\langle \mathbf{p}_2 | S | \mathbf{p}_1 \rangle$ for some $\mathbf{p}_2, \mathbf{p}_1 \in \mathbb{R}^3$. It is equal to

$$\langle \mathbf{p}_2 | 1 | \mathbf{p}_1 \rangle + (-ig) \int_{-\infty}^{\infty} dt \langle \mathbf{p}_2 | e^{itH_0} V e^{-itH_0} | \mathbf{p}_1 \rangle + \mathcal{O}(g^2).$$

Recall that

$$\langle \mathbf{p}_2 | 1 | \mathbf{p}_1 \rangle = \langle \mathbf{p}_2 | \mathbf{p}_1 \rangle = (2\pi)^3 \delta^{(3)}(\mathbf{p}_2 - \mathbf{p}_1).$$

Next, recall that as a function on momentum space, the state $|\mathbf{p}_1\rangle$ is the function

$$\psi(\mathbf{p}) = (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{p}_1).$$

Since H_0 on momentum space is just multiplication by $\mathbf{p}^2/2m$,

$$\begin{aligned}
(e^{-itH_0}\psi)(\mathbf{p}) &= e^{-it\mathbf{p}^2/2m}\psi(\mathbf{p}) \\
&= e^{-it\mathbf{p}^2/2m}(2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{p}_1) \\
&= e^{-it\mathbf{p}_1^2/2m}(2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{p}_1) = e^{-it\mathbf{p}_1^2/2m}\psi(\mathbf{p}).
\end{aligned}$$

In short,

$$e^{-itH_0}|\mathbf{p}_1\rangle = e^{-it\mathbf{p}_1^2/2m}|\mathbf{p}_1\rangle.$$

Carrying out a similar calculation for $\mathbf{p}_2$, we get

$$\langle \mathbf{p}_2 | e^{itH_0} V e^{-itH_0} | \mathbf{p}_1 \rangle = e^{it(\mathbf{p}_2^2 - \mathbf{p}_1^2)/2m} \langle \mathbf{p}_2 | V | \mathbf{p}_1 \rangle.$$

Next, recall that in position space, $|\mathbf{p}\rangle$ is represented by the function $\varphi(\mathbf{x}) = e^{i\mathbf{x}\cdot\mathbf{p}}$. Therefore a position space calculation gives

$$\langle \mathbf{p}_2 | V | \mathbf{p}_1 \rangle = \int d^3\mathbf{x} e^{-i\mathbf{x}\cdot(\mathbf{p}_2 - \mathbf{p}_1)} V(\mathbf{x}) = \widehat{V}(\mathbf{p}_2 - \mathbf{p}_1).$$

Combining, we get

$$\begin{aligned}
&\langle \mathbf{p}_2 | S | \mathbf{p}_1 \rangle \\
&= (2\pi)^3 \delta^{(3)}(\mathbf{p}_2 - \mathbf{p}_1) + (-ig) \int_{-\infty}^{\infty} dt e^{it(\mathbf{p}_2^2 - \mathbf{p}_1^2)/2m} \widehat{V}(\mathbf{p}_2 - \mathbf{p}_1) + \mathcal{O}(g^2) \\
&= (2\pi)^3 \delta^{(3)}(\mathbf{p}_2 - \mathbf{p}_1) + (-ig)(2\pi) \widehat{V}(\mathbf{p}_2 - \mathbf{p}_1) \delta((\mathbf{p}_2^2 - \mathbf{p}_1^2)/2m) + \mathcal{O}(g^2).
\end{aligned}$$

This is called the *Born approximation*. Note that the first delta function is a delta function in $\mathbb{R}^3$ and the second one is a delta function in $\mathbb{R}$.

15.2. What does it mean?

The meaning of the Born approximation is not transparent from the representation in terms of delta functions. Let us now try to understand it better. Suppose that the incoming state was a proper state $|\psi\rangle$ instead of the improper state $|\mathbf{p}_1\rangle$. Then $f(\mathbf{p}_2) = |\langle \mathbf{p}_2 | S | \psi \rangle|^2$ would be proportional to the probability density of the momentum of the outgoing state. Let us approximate the improper state $|\mathbf{p}_1\rangle$ by the proper state $|\psi_\varepsilon\rangle$, represented in momentum space by the Gaussian density

$$\psi_\varepsilon(\mathbf{p}) = \frac{1}{(2\pi)^{3/2}\varepsilon^{3/2}} e^{-(\mathbf{p}-\mathbf{p}_1)^2/2\varepsilon},$$

where ε is some small number. With this approximation, the 3D delta function $\delta^{(3)}(\mathbf{p}_2 - \mathbf{p}_1)$ in the Born approximation gets replaced by something like

$$\frac{C_1}{\varepsilon^{3/2}} \exp\left(-\frac{(\mathbf{p}_2 - \mathbf{p}_1)^2}{2\varepsilon} \right),$$

where C_1 is a constant. In a similar way, the one-dimensional delta function $\delta((\mathbf{p}_2^2 - \mathbf{p}_1^2)/2m)$ gets replaced by something like

$$\frac{C_2}{\sqrt{\varepsilon}} \exp\left(-\frac{(\mathbf{p}_2^2 - \mathbf{p}_1^2)^2}{8m^2\varepsilon} \right).$$

Then, we have $f(\mathbf{p}_2) = |\langle \mathbf{p}_2 | S | \psi_\varepsilon \rangle|^2 = |A(\mathbf{p}_2) + B(\mathbf{p}_2) + \mathcal{O}(g^2)|^2$ with:

$$A(\mathbf{p}_2) = \frac{C_1 (2\pi)^3}{\varepsilon^{3/2}} \exp\left(-\frac{(\mathbf{p}_2 - \mathbf{p}_1)^2}{2\varepsilon} \right)$$

and

$$B(\mathbf{p}_2) = C_2 (-ig)(2\pi) \widehat{V}(\mathbf{p}_2 - \mathbf{p}_1) \frac{1}{\sqrt{\varepsilon}} \exp\left(-\frac{(\mathbf{p}_2^2 - \mathbf{p}_1^2)^2}{8m^2\varepsilon} \right).$$

Let S be the sphere of center $\mathbf{0}$ and radius $|\mathbf{p}_1|$ in $\mathbb{R}^3$. Let $\mathcal{B}_\varepsilon$ denote the ball of radius $\sqrt{\varepsilon}$ around $\mathbf{p}_1$. Let $\mathcal{A}_\varepsilon$ denote the annulus of width $\sqrt{\varepsilon}$ around the sphere S. Then, roughly speaking,

- $A(\mathbf{p}_2)$ is of order $\varepsilon^{-3/2}$ in $\mathcal{B}_\varepsilon$, and very small outside.
- $B(\mathbf{p}_2)$ is of order $\varepsilon^{-1/2}$ in $\mathcal{A}_\varepsilon$, and very small outside.

This shows that, again roughly speaking,

- $f(\mathbf{p}_2)$ is of order ε^{-3} in $\mathcal{B}_\varepsilon$ and very small outside.
- $f(\mathbf{p}_2)$ is of order ε^{-1} in $\mathcal{A}_\varepsilon$ and very small outside.

Moreover, the volume of $\mathcal{B}_\varepsilon$ is of order $\varepsilon^{3/2}$, and the volume of $\mathcal{A}_\varepsilon$ is of order $\varepsilon^{1/2}$. Thus, the integral of $f(\mathbf{p}_2)$ in $\mathcal{B}_\varepsilon$ is of order $\varepsilon^{-3} \cdot \varepsilon^{3/2} = \varepsilon^{-3/2}$, and the integral of $f(\mathbf{p}_2)$ in $\mathcal{A}_\varepsilon$ is of order $\varepsilon^{-1} \cdot \varepsilon^{1/2} = \varepsilon^{-1/2}$. Everywhere else, the integral is negligible. Since $\varepsilon^{-3/2} \gg \varepsilon^{-1/2}$, this shows that as $\varepsilon \to 0$, the probability density fully concentrates near $\mathbf{p}_1$. Thus, if a particle comes in with momentum $\approx \mathbf{p}_1$, it also goes out with momentum $\approx \mathbf{p}_1$.

This looks uninteresting, but the interesting thing happens if we *condition* on the event that the particle scatters. When the incoming state is $|\psi_\varepsilon\rangle$, the above calculations show that the conditional density of the outgoing momentum given $|\mathbf{p}_2 - \mathbf{p}_1| > \eta$ for η small is approximately proportional to $|\widehat{V}(\mathbf{p}_2 - \mathbf{p}_1)|^2 + \mathcal{O}(g)$ in $\mathcal{A}_\varepsilon$.

If we fix η and send $\varepsilon \to 0$, and then send $\eta \to 0$, we get the probability distribution of the outgoing momentum *conditional on it being different than the incoming momentum.* This probability distribution is supported on S, with probability density proportional to $|\widehat{V}(\mathbf{p}_2 - \mathbf{p}_1)|^2 + \mathcal{O}(g)$.

To summarize, the Born approximation says that if a particle comes in with momentum exactly equal to $\mathbf{p}_1$, it also goes out with momentum exactly equal to $\mathbf{p}_1$, but conditional on the probability zero event that the outgoing momentum $\neq \mathbf{p}_1$, the outgoing momentum has a probability distribution supported on the sphere of radius $|\mathbf{p}_1|$ and center $\mathbf{0}$ in $\mathbb{R}^3$, with density proportional to $f(\mathbf{p}_2) = |\widehat{V}(\mathbf{p}_2 - \mathbf{p}_1)|^2 + \mathcal{O}(g)$.

It may seem strange to condition on a probability zero event, but this is routinely done in probability theory, for example conditioning on the endpoint of Brownian motion being zero to get the Brownian bridge. Moreover, since it is impossible to generate quantum particles with momentum exactly equal to a given value, the above statement is only an idealization. In practical terms, it means that if the momentum of an incoming particle is very close to a given value with high probability, then the outgoing momentum is also close to same value with high probability, but if by chance it scatters, then the Born approximation gives a first order approximation for the probability amplitude of the outgoing momentum. This probability distribution is supported on the set of momenta that have the same Euclidean norm as the incoming momentum, in agreement with conservation of energy.

LECTURE 16

Hamiltonian densities

Date: 10/29/2018
Scribe: Jaydeep Singh

16.1. Scattering In QFT

Recall the setting for QFT, as applied to scalar bosons of mass m. Here 'scalar' implies the underlying Hilbert space of the single-particle system is

$$\mathcal{H} = L^2(X_m, d\lambda_m),\ X_m = \{p \in \mathbb{R}^{1,3} : p^2 = m^2,\ p^0 \geq 0\}.$$

Here X_m is the mass shell, $d\lambda_m$ is the Lorentz invariant measure, and $p^2 = (p,p) = (p^0)^2 - (p^1)^2 - (p^2)^2 - (p^3)^2$ is the Minkowski norm of p. From $\mathcal{H}$ we constructed the multi-particle Bosonic Fock space $\mathcal{B}$, and defined the free evolution Hamiltonian H_0. The action of the latter on the state $\psi(p_1, \ldots, p_n)$ is given by

$$(H_0\psi)(p_1, \ldots, p_n) = \left(\sum_{j=1}^{n} (p_j^0)^2\right)\psi(p_1, \ldots, p_n).$$

We further defined creation and annihilation operators as operator-valued distributions $a^\dagger(p), a(p)$ on X_m, and variants $a(\mathbf{p})$ and $a^\dagger(\mathbf{p})$. The relation between these distributions is given by

$$a^\dagger(\mathbf{p}) = \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} a^\dagger(p), \quad a(\mathbf{p}) = \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} a(p),$$

where $\omega_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m^2}$. Finally, recall the massive scalar free field:

$$\varphi(x) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \left(e^{-i(x,p)} a(\mathbf{p}) + e^{i(x,p)} a^\dagger(\mathbf{p})\right).$$

It will be useful to define

$$\varphi^-(x) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} e^{i(x,p)} a^\dagger(\mathbf{p}),\ \varphi^+(x) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} e^{-i(x,p)} a(\mathbf{p}),$$

where $\varphi^-(x)$ and $\varphi^+(x)$ are called creation and annihilation parts respectively.

Suppose we have the Hamiltonian $H = H_0 + gH_I$, with H_I an interaction Hamiltonian, g a coupling constant, and S the associated scattering operator. As in non-relativistic QM, we aim to understand the amplitudes of

scattering processes for general Hamiltonians, which reduces to understanding the scattering operator S. In QFT, however, we must allow for particle number to change: For $\mathbf{p}_1, \dots, \mathbf{p}_k$ incoming momenta and $\mathbf{q}_1, \dots, \mathbf{q}_j$ outgoing momenta, we wish to compute the amplitudes

$$\langle \mathbf{q}_1, \dots, \mathbf{q}_j | S | \mathbf{p}_1, \dots, \mathbf{p}_k \rangle.$$

The Dyson series expansion remains valid in the QFT setting, since the formal derivation is exactly the same. When H_I is constructed using objects known as Hamiltonian densities, and the Hamiltonian densities satisfy certain properties, the Dyson series expansion takes a particularly nice form. We will work this out in this lecture.

16.2. Construction of interaction Hamiltonians

We will now see a general prescription for constructing interaction Hamiltonians. Due to the difficulties associated with particle creation and annihilation, we will be vague about the underlying Hilbert space on which these operators will act.

DEFINITION 16.1. A **Hamiltonian density** is an operator-valued distribution, with kernel denoted both as $\mathcal{H}(x)$ and $\mathcal{H}(t, \mathbf{x})$ for $x = (t, \mathbf{x}) \in \mathbb{R}^{1,3}$. This distribution must satisfy the following:

(1) (Time evolution.) $\mathcal{H}(t, \mathbf{x}) = e^{itH_0} \mathcal{H}(0, \mathbf{x}) e^{-itH_0}$.
(2) (Equal time commutation.) For any $t \in \mathbb{R}$ and any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$, $[\mathcal{H}(t, \mathbf{x}), \mathcal{H}(t, \mathbf{y})] = 0$.

The associated **interaction Hamiltonian** H_I is the operator-valued distribution

$$H_I = \int_{\mathbb{R}^3} d^3\mathbf{x} \mathcal{H}(0, \mathbf{x}).$$

REMARK 16.1. In a physical setting, one imposes the following additional constraints:

(1) The distribution $\mathcal{H}(x)$ should be Lorentz invariant.
(2) $\mathcal{H}(x), \mathcal{H}(y)$ should commute whenever x and y are spacelike separated, i.e. $(x - y)^2 < 0$.

We do not need to verify these conditions for our main purpose, which is to get a certain form of the Dyson expansion.

We now turn to an example of a Hamiltonian constructed via a density, namely the φ^4 theory. Here, the Hamiltonian density is

$$\mathcal{H}(x) = \frac{1}{4!} :\varphi(x)^4:.$$

In the remainder of this section, we prove $\mathcal{H}(x)$ satisfies conditions (1) and (2) of Definition 16.1.

LEMMA 16.1. *For all $k \in \mathbb{N}$, $:\varphi(x)^k:$ is a formal polynomial in $\varphi(x)$.*

PROOF. The proof is by induction on k. The $k = 1$ case follows immediately, as Wick ordering has no impact on the expression for $\varphi(x)$. So suppose that the result is given for $k \leq n$. First, writing $\varphi(x) = \varphi^-(x) + \varphi^+(x)$, we observe that after Wick ordering, the product $(\varphi^-(x) + \varphi^+(x))^n$ collapses as though the $\varphi^\pm$ commute:

$$:\varphi(x)^n: = :(\varphi^-(x) + \varphi^+(x))^n: = \sum_{j=0}^{n} \binom{n}{j} \varphi^-(x)^j \varphi^+(x)^{n-j}.$$

Next, we turn to the commutator $[\varphi^-(x), \varphi^+(x)]$:

$$[\varphi^+(x), \varphi^-(x)] = \int\int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{d^3\mathbf{q}}{(2\pi)^3} \frac{1}{\sqrt{2\omega_\mathbf{p}}} \frac{1}{\sqrt{2\omega_\mathbf{q}}} e^{i(x,q-p)} [a(\mathbf{p}), a^\dagger(\mathbf{q})].$$

But we have seen earlier that the commutator in the integrand is just $(2\pi)^3 \delta^{(3)}(\mathbf{q} - \mathbf{p})1$, so integration over $d^3\mathbf{q}$ fixes the value of $\mathbf{q}$. Since $q = (\omega_\mathbf{q}, \mathbf{q})$, we see q is only a function of $\mathbf{q}$, and so integrating the delta function in fact sets $q = p$. This procedure thus yields

$$[\varphi^+(x), \varphi^-(x)] = \left(\int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{2\omega_\mathbf{p}} \right) 1.$$

Let us denote the term within the brackets by C. Note C is not a finite quantity, but just a symbol in our formal calculations.

We now employ the commutation relation repeatedly to compute

$$\begin{aligned}
&\varphi^-(x)^j \varphi^+(x)^{n-j} \varphi^-(x) \\
&= \varphi^-(x)^j \varphi^+(x)^{n-j-1} \varphi^-(x) \varphi^+(x) + C\varphi^-(x)^j \varphi^+(x)^{n-j-1} \\
&= \varphi^-(x)^j \varphi^+(x)^{n-j-2} \varphi^-(x) \varphi^+(x)^2 + 2C\varphi^-(x)^j \varphi^+(x)^{n-j-1} \\
&= \cdots \\
&= \varphi^-(x)^{j+1} \varphi^+(x)^{n-j} + (n-j) C \varphi^-(x)^j \varphi^+(x)^{n-j-1}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
:\varphi(x)^n:\varphi(x) &= \left(\sum_{j=0}^{n} \binom{n}{j} \varphi^-(x)^j \varphi^+(x)^{n-j} \right) \left(\varphi^+(x) + \varphi^-(x) \right) \\
&= \sum_{j=0}^{n} \binom{n}{j} \varphi^-(x)^j \varphi^+(x)^{n-j+1} + \sum_{j=0}^{n} \binom{n}{j} \varphi^-(x)^j \varphi^+(x)^{n-j} \varphi^-(x) \\
&= \sum_{j=0}^{n+1} \binom{n+1}{j} \varphi^-(x)^j \varphi^+(x)^{n+1-j} + \sum_{j=0}^{n-1} C(n-j) \binom{n}{j} \varphi^-(x)^j \varphi^+(x)^{n-j-1} \\
&= \sum_{j=0}^{n+1} \binom{n+1}{j} \varphi^-(x)^j \varphi^+(x)^{n+1-j} + \sum_{i=0}^{n-1} Cn \binom{n-1}{j} \varphi^-(x)^j \varphi^+(x)^{n-1-j},
\end{aligned}$$

where we have employed the previous display and the identity

$$\binom{n}{j} + \binom{n}{j-1} = \binom{n+1}{j}$$

in the second-to-last equality, and the identity

$$(n-j)\binom{n}{j} = n\binom{n-1}{j}$$

in the final step. Thus, we get

$$:\varphi(x)^n:\varphi(x) = :\varphi^{n+1}(x): + Cn\varphi^{n-1}(x):,$$

which clearly completes the induction step. □

PROPOSITION 16.1. *The Hamiltonian density* $\mathcal{H}(x) = :\varphi(x)^k:$ *satisfies conditions* (1) *and* (2) *of Definition* 16.1.

PROOF. In addition to the lemma above, we will need the following facts:

(1) For all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$, $[\varphi(0,\mathbf{x}), \varphi(0,\mathbf{y})] = 0$. This follows directly from a computation analogous to the one given above.
(2) The time evolution of the free field satisfies

$$\varphi(t,\mathbf{x}) = e^{itH_0}\varphi(0,\mathbf{x})e^{-itH_0},$$

which was observed in an earlier lecture.

We may generalize the second fact to arbitrary polynomials of φ, using

$$e^{itH_0}\varphi(0,\mathbf{x})^k e^{-itH_0} = (e^{itH_0}\varphi(0,\mathbf{x})e^{-itH_0})^k = \varphi(t,\mathbf{x})^k.$$

Thus, using that $:\varphi(t,\mathbf{x})^k: = f(\varphi(t,\mathbf{x}))$ for f a formal polynomial, we conclude

$$:\varphi(t,\mathbf{x})^k: = f(\varphi(t,\mathbf{x})) = e^{itH_0}f(\varphi(0,\mathbf{x}))e^{-itH_0} = e^{itH_0}:\varphi(0,\mathbf{x})^k:e^{-itH_0},$$

which proves the time evolution property. Similarly, using the first fact above, and the observation that A, B commute implies A^m, B^n commute, we get

$$\begin{aligned}[\varphi(t,\mathbf{x})^m, \varphi(t,\mathbf{y})^n] &= [e^{itH_0}\varphi(0,\mathbf{x})^m e^{-itH_0}, e^{itH_0}\varphi(0,\mathbf{y})^n e^{-itH_0}] \\ &= e^{itH_0}[\varphi(0,\mathbf{x})^m, \varphi(0,\mathbf{y})^n]e^{-itH_0} = 0.\end{aligned}$$

Thus, again writing $:\varphi(t,\mathbf{x})^k: = f(\varphi(t,\mathbf{x}))$ for some formal polynomial f, and exploiting the bilinearity of the commutator, we see that

$$[:\varphi(t,\mathbf{x})^k:, :\varphi(t,\mathbf{y})^k:]$$

is a sum of commutators of the form $[\varphi(t,\mathbf{x})^m, \varphi(t,\mathbf{y})^n]$, all of which vanish by the above argument. Thus the equal time commutation property holds for $\mathcal{H}$. □

16.3. Dyson series for the QFT scattering operator

Finally, we derive a simplified formula for the Dyson series representation in QFT, assuming that the interaction Hamiltonian arises from a Hamiltonian density subject to the conditions (1) and (2) of Definition 16.1. We first observe the following:

(1) Applying the time evolution of the interaction Hamiltonian, we have

$$\begin{aligned} H_I(t) &= e^{itH_0} H_I(0) e^{-itH_0} \\ &= e^{itH_0} \left(\int d^3\mathbf{x} \mathcal{H}(0,\mathbf{x}) \right) e^{-itH_0} \\ &= \int d^3\mathbf{x} e^{itH_0} \mathcal{H}(0,\mathbf{x}) e^{-itH_0} \\ &= \int d^3\mathbf{x} \mathcal{H}(t,\mathbf{x}). \end{aligned}$$

The interaction Hamiltonian at time t is thus naturally determined by the Hamiltonian density at time t.

(2) The equal time commutation property ensures that the time ordered product $\mathcal{T}\mathcal{H}(x_1)\cdots\mathcal{H}(x_n)$ is unambiguously defined for any $x_1, \ldots, x_n \in \mathbb{R}^{1,3}$.

So, inserting $H_I(t) = \int d\mathbf{x}^3 \mathcal{H}(t,\mathbf{x})$ into the formula for the nth term S_n in the Dyson series expansion, we observe that each argument is integrated over all of spacetime. We may thus replace all integrals with integrals over $\mathbb{R}^{1,3}$, yielding

$$S_n = \frac{(-ig)^n}{n!} \int_{\mathbb{R}^{1,3}} \cdots \int_{\mathbb{R}^{1,3}} dx_1 \cdots dx_n \mathcal{T}\mathcal{H}(x_1)\cdots\mathcal{H}(x_n).$$

This representation of the Dyson series will be useful going forward.

LECTURE 17

Wick's theorem

Date: 10/31/2018
Scribe: Sungyeon Yang

17.1. Calculating the Dyson series: First steps

In this lecture we begin the process of learning how to compute the terms in the Dyson series expansion for φ^4 theory. Recall that the Hilbert space of interest is $\mathcal{H} = L^2(X_m, d\lambda_m)$ and we have operator-valued distributions $A, A^\dagger$ acting on this space. Let $\mathcal{C}$ be the class of all operators on $\mathcal{B}_0$ of the form $A(\xi) + A^\dagger(\eta)$ for some $\xi, \eta \in \mathcal{H}$.

LEMMA 17.1. *If $B_1 = A(\xi_1) + A^\dagger(\eta_1)$ and $B_2 = A(\xi_2) + A^\dagger(\eta_2)$, then we have*

$$\langle 0|B_1 B_2|0\rangle = (\xi_1, \eta_2).$$

PROOF. It is easy to see from the definitions of A and $A^\dagger$ that for any ξ, $A(\xi)|0\rangle = 0$ and $\langle 0|A^\dagger(\xi) = 0$. Thus,

$$\begin{aligned}\langle 0|B_1 B_2|0\rangle &= \langle 0|(A(\xi_1) + A^\dagger(\eta_1))(A(\xi_2) + A^\dagger(\eta_2))|0\rangle \\ &= \langle 0|A(\xi_1)A^\dagger(\eta_2)|0\rangle.\end{aligned}$$

The proof is now easily completed using the commutation relation

$$[A(\xi), A^\dagger(\eta)] = (\xi, \eta)1$$

that we derived earlier. □

Note that $a(\mathbf{p})$, $a^\dagger(\mathbf{p})$ and $\varphi(x)$ are elements of $\mathcal{C}$ as

$$a(\mathbf{p}) = \frac{a(p)}{\sqrt{2\omega_{\mathbf{p}}}} = \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} A(\delta_p), \quad a^\dagger(\mathbf{p}) = \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} A^\dagger(\delta_p),$$

and

$$\varphi(x) = A(f_x) + A^\dagger(f_x),$$

where $f_x = e^{i(x,p)}$. Here, as usual, $p = (\omega_{\mathbf{p}}, \mathbf{p})$. For the last claim, note that

$$\begin{aligned}\varphi(x) &= \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\mathbf{p}}}}(e^{-i(x,p)}a(\mathbf{p}) + e^{i(x,p)}a^\dagger(\mathbf{p})) \\ &= \int \frac{d^3\mathbf{p}}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{p}}}(e^{-i(x,p)}a(p) + e^{i(x,p)}a^\dagger(p)) \\ &= \int d\lambda_m(p)(e^{-i(x,p)}a(p) + e^{i(x,p)}a^\dagger(p)) \\ &= A(f_x) + A^\dagger(f_x).\end{aligned}$$

Consider the amplitude $\langle 0|a(\mathbf{p}')a^\dagger(\mathbf{p})|0\rangle$. By the commutation relation for $a(\mathbf{p})$ and $a^\dagger(\mathbf{p})$,

$$a(\mathbf{p}')a^\dagger(\mathbf{p}) = a^\dagger(\mathbf{p})a(\mathbf{p}') + (2\pi)^3\delta^{(3)}(\mathbf{p} - \mathbf{p}')1.$$

On the other hand, $a(\mathbf{p}')\,|0\rangle = 0$ and $\langle 0|\,a^\dagger(\mathbf{p}) = 0$. Combining, we get

$$\langle 0|a(\mathbf{p}')a^\dagger(\mathbf{p})|0\rangle = (2\pi)^3\delta^{(3)}(\mathbf{p} - \mathbf{p}').$$

We also have

$$\langle 0|a^\dagger(\mathbf{p}')a(\mathbf{p})|0\rangle = 0, \quad \langle 0|a(\mathbf{p}')a(\mathbf{p})|0\rangle = 0, \quad \langle 0|a^\dagger(\mathbf{p}')a^\dagger(\mathbf{p})|0\rangle = 0.$$

By Lemma 17.1 and the observations made following the proof of the lemma, we have

$$\begin{aligned}\langle 0|a(\mathbf{p})\varphi(x)|0\rangle &= \frac{1}{\sqrt{2\omega_{\mathbf{p}}}}(\delta_p, f_x) = \frac{1}{\sqrt{2\omega_{\mathbf{p}}}}e^{i(x,p)}, \\ \langle 0|\varphi(x)a^\dagger(\mathbf{p})|0\rangle &= \frac{1}{\sqrt{2\omega_{\mathbf{p}}}}(f_x, \delta_p) = \frac{1}{\sqrt{2\omega_{\mathbf{p}}}}e^{-i(x,p)}.\end{aligned}$$

We will need these computations later.

17.2. Wick's theorem

Let us now introduce the main tool for computing the terms in Dyson's expansion in QFT. If k is an even number, and pairing l of k is a permutation $(l_1, l_1', l_2, l_2', \ldots, l_{k/2}, l_{k/2}')$ of $(1, 2, \ldots, k)$ such that $l_j < l_j'$ for all j. The following result is known as Wick's theorem.

THEOREM 17.1. *If $B_1, \ldots, B_k \in \mathcal{C}$, then we have*

$$\langle 0|B_1B_2\cdots B_k|0\rangle = \sum_{\text{pairings } l}\left(\prod_{j=1}^{k/2}\langle 0|B_{l_j}B_{l_j'}|0\rangle\right)$$

for k even, and $\langle 0|B_1B_2\cdots B_k|0\rangle = 0$ for k odd.

PROOF. We prove the theorem by induction on k. Let $B_j = A(\xi_j) + A^\dagger(\eta_j)$. Then,

$$\begin{aligned}\langle 0|B_1 \cdots B_{k-1}B_k|0\rangle &= \langle 0|B_1 \cdots B_{k-1}(A(\xi_k) + A^\dagger(\eta_k))|0\rangle \\ &= \langle 0|B_1 \cdots B_{k-1}A^\dagger(\eta_k)|0\rangle,\end{aligned}$$

since $A(\xi_k)\,|0\rangle = 0$. Since $[A^\dagger(\xi), A^\dagger(\eta)] = 0$ for any ξ and η, Lemma 17.1 gives

$$[B_{k-1}, A^\dagger(\eta_k)] = [A(\xi_{k-1}), A^\dagger(\eta_k)] = (\xi_{k-1}, \eta_k)1.$$

Thus,

$$\begin{aligned}\langle 0|B_1 \cdots B_{k-1}A^\dagger(\eta_k)|0\rangle = \langle 0|B_1 \cdots B_{k-2}A^\dagger(\eta_k)B_{k-1}|0\rangle \\ + (\xi_{k-1}, \eta_k)\langle 0|B_1 \cdots B_{k-2}|0\rangle.\end{aligned}$$

We can iterate this step until $A^\dagger(\eta_k)$ moves all the way to the left, which gives zero since $\langle 0|A^\dagger(\eta_k) = 0$. After the final step, we get

$$\langle 0|B_1B_2 \cdots B_k|0\rangle = \sum_{j=1}^{k-1}(\xi_j, \eta_k)\langle 0|B_1 \cdots \widehat{B}_j \cdots B_{k-1}|0\rangle$$

where the hatted term is omitted. The induction step is now easily completed by recalling that $(\xi_j, \eta_k) = \langle 0|B_jB_k|0\rangle$. □

The number of terms in the sum in Wick's theorem is a well-known quantity. It is equal to

$$(k-1)!! := (k-1)(k-3)\cdots 5 \cdot 3 \cdot 1.$$

When applying Wick's theorem, verifying that the total number of terms considered is indeed $(k-1)!!$ is one way of ensuring that we have not missed out anything.

17.3. Contraction diagrams

Each $\langle 0|B_jB_k|0\rangle$ in Wick's theorem is called a *contraction* of B_j and B_k. The sum in Wick's theorem is convenient to handle using contraction diagrams. Diagramatically, we represent each B_j by a vertex, with an edge hanging out. Then we tie up each tail with one other, so that there is no untied tail. Each such diagram contributes one term to the sum in Wick's theorem.

Consider, for example, the calculation of $\langle 0|B_1B_2B_3B_4|0\rangle$. The vertices with freely hanging edges, and the three diagrams obtained by tying up the edges, are shown in Figure 17.1.

When some B_i occurs as a power B_i^k, then it is represented as a single vertex with k distinct tails hanging out. Consider $\langle 0|B_1B_2^2B_3|0\rangle$ for example. There are three diagrams in this calculation, but one is repeated twice. So the pictorial representation shows two diagrams. This is shown in Figure 17.2.

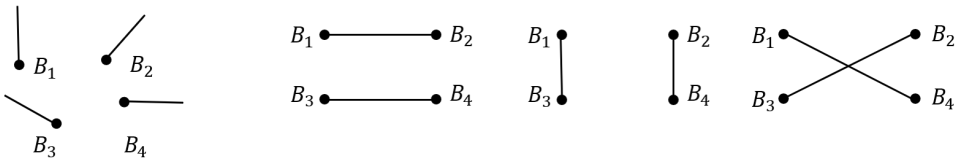

FIGURE 17.1. Diagrams for $\langle 0|B_1B_2B_3B_4|0\rangle$. The vertices with freely hanging edges are shown on the left.

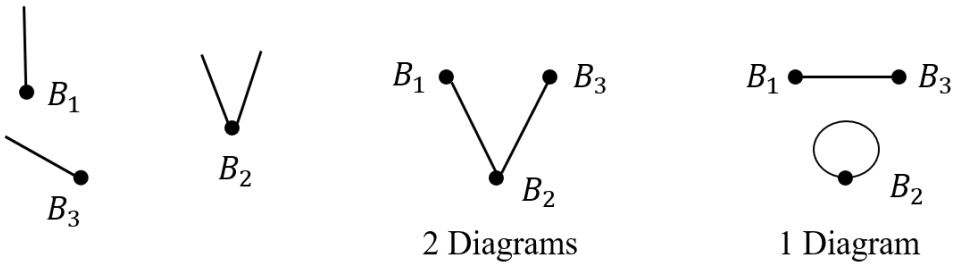

FIGURE 17.2. Diagrams for $\langle 0|B_1B_2^2B_3|0\rangle$. The vertices with freely hanging edges are shown on the left. The first diagram is repeated twice because the two edges hanging out from B_2 are distinct.

As a final example, consider $\langle 0|B_1B_2B_3^4B_4B_5|0\rangle$. Since there are 8 operators, there should be $(8-1)!! = 105$ diagrams. Fortunately, there are many repeated 'similar' diagrams, which makes it easy to represent them compactly. This is shown in Figure 17.3.

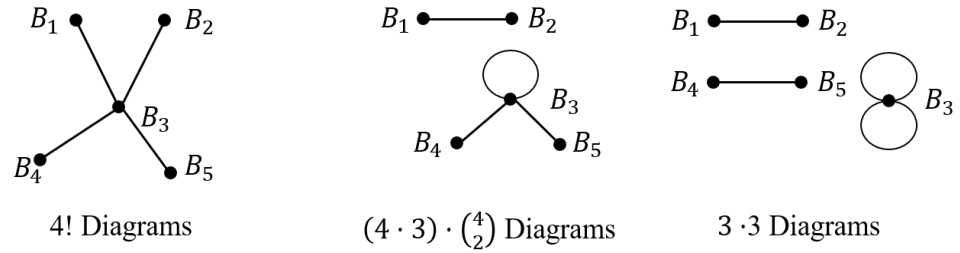

FIGURE 17.3. Diagrams for $\langle 0|B_1B_2B_3^4B_4B_5|0\rangle$. For the second case, there are 12 diagrams after fixing two vertices, and 6 ways of choosing the two vertices. There are $24 + 72 + 9 = 105 = (8-1)!!$ diagrams in total.

LECTURE 18

A first-order calculation in φ^4 theory

Date: 11/2/2018
Scribe: Jing An

18.1. Contraction diagrams under normal ordering

Last time, we investigated the calculations and diagrams for

$$\langle 0|B_1 \cdots B_i^k \cdots B_m|0\rangle. \tag{18.1}$$

We want to understand what happens if B_i^k is replaced by $:B_i^k:$.

PROPOSITION 18.1. *If we replace B_i^k by $:B_i^k:$ in* (18.1)*, then the answer is obtained by deleting all diagrams in the original expansion that have any self-loops at B_i.*

For instance, recall the example $\langle 0|B_1B_2B_3^4B_4B_5|0\rangle$, whose diagrams were displayed in Figure 17.3 in the previous lecture. If we replace B_3^4 by $:B_3^4:$, then only the first set of 4! diagrams will survive.

PROOF OF PROPOSITION 18.1. Since $B_i \in \mathcal{C}$, it can be written in the form $B_i = A(\xi_i) + A^\dagger(\eta_i)$ for some $\xi_i, \eta_i \in \mathcal{H}$. Let us denote $B_i^+ := A(\xi_i)$ and $B_i^- := A^\dagger(\eta_i)$ for notational simplicity. Then

$$B_i^k = (B_i^+ + B_i^-)^k.$$

Expanding it, we get $\langle 0|B_1 \cdots B_i^k \cdots B_m|0\rangle$ = sum of 2^k such quantities. Similarly, because

$$:B_i^k: = \sum_{j=0}^{k} \binom{k}{j} (B_i^-)^j (B_i^+)^{k-j},$$

we will again get a set of 2^k terms. Moreover, there is a natural correspondence between these two sets of 2^k terms. We take any such term X on one side and the corresponding Y on the other side, and write down the diagrams for X and Y. Note that each diagram gives a product of contractions. If a contraction for X is of the form $\langle 0|B_jB_i^+|0\rangle$ for some $j < i$ or $\langle 0|B_i^+B_j|0\rangle$ for some $j > i$, then it is the same in both diagrams, and likewise for B_i^-. Moreover,

- Any contraction like $\langle 0|B_i^+B_i^+|0\rangle$ or $\langle 0|B_i^-B_i^-|0\rangle$ is the same in both diagrams and actually it is equal to 0, because $B_i^+|0\rangle = 0$ and $\langle 0|B_i^- = 0$.

- Any contraction like $\langle 0|B_i^- B_i^+|0\rangle$ is the same in both diagrams and is equal to 0 due to the same reason as above.
- Any contraction like $\langle 0|B_i^+ B_i^-|0\rangle$ (which is nonzero) in a diagram for X has to be replaced by $\langle 0|B_i^- B_i^+|0\rangle$ (which is 0) in the corresponding diagram for Y.

This shows that the diagram for Y can be computed by taking the diagram for X and replacing it by 0 if there exists any contraction like

$$\langle 0|B_i^+ B_i^+|0\rangle, \langle 0|B_i^- B_i^-|0\rangle, \langle 0|B_i^- B_i^+|0\rangle, \text{ or } \langle 0|B_i^+ B_i^-|0\rangle$$

in the diagram. (The first three above are automatically 0 and the last one, by Wick's ordering, can be replaced by $\langle 0|B_i^- B_i^+|0\rangle$.) But note that such terms arise from terms like $\langle 0|B_i B_i|0\rangle$ in the original Wick expansion for $\langle 0|B_1 \cdots B_i^k \cdots B_m|0\rangle$.

This tells us that if we take the Wick expansion for $\langle 0|B_1 \cdots B_i^k \cdots B_m|0\rangle$ and remove any diagram that has a contraction of the form $\langle 0|B_i B_i|0\rangle$, then we will get the Wick expansion for $\langle 0|B_1 \cdots :B_i^k: \cdots B_m|0\rangle$. □

18.2. A first-order calculation in φ^4 theory

Consider φ^4 theory. Suppose we have distinct $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4$, and we want to compute

$$\langle \mathbf{p}_3, \mathbf{p}_4|S|\mathbf{p}_1, \mathbf{p}_2\rangle$$

to the first order in perturbation theory. Recall the Dyson series expansion

$$\begin{aligned} S &= 1 + \sum_{n=1}^{\infty} \frac{(-ig)^n}{n!} \int \cdots \int d^4x_1 \cdots d^4x_n \mathcal{T}\left(\frac{1}{4!}:\varphi(x_1)^4: \cdots \frac{1}{4!}:\varphi(x_n)^4:\right) \\ &= 1 - \frac{ig}{4!} \int d^4x \, :\varphi(x)^4: + \mathcal{O}(g^2). \end{aligned}$$

Since $\mathbf{p}_i$'s are distinct, we have $\langle \mathbf{p}_3, \mathbf{p}_4|\mathbf{p}_1, \mathbf{p}_2\rangle = 0$. Therefore,

$$\langle \mathbf{p}_3, \mathbf{p}_4|S|\mathbf{p}_1, \mathbf{p}_2\rangle = -\frac{ig}{4!} \int d^4x \langle \mathbf{p}_3, \mathbf{p}_4|:\varphi(x)^4:|\mathbf{p}_1, \mathbf{p}_2\rangle + \mathcal{O}(g^2).$$

Note that

$$\langle \mathbf{p}_3, \mathbf{p}_4|:\varphi(x)^4:|\mathbf{p}_1, \mathbf{p}_2\rangle = \langle 0|a(\mathbf{p}_3)a(\mathbf{p}_4):\varphi(x)^4:a^\dagger(\mathbf{p}_1)a^\dagger(\mathbf{p}_2)|0\rangle. \tag{18.2}$$

The set of contraction diagrams for the above quantity consists of 4! diagrams like the one displayed in Figure 18.1.

Feynman diagrams are contraction diagrams but without labels and with arrows denoting incoming and outgoing particles. The contraction diagram of Figure 18.1 becomes the Feynman diagram of Figure 18.2. The usual convention for Feynman diagrams is that particles are shown to be coming in from the left and exiting on the right.

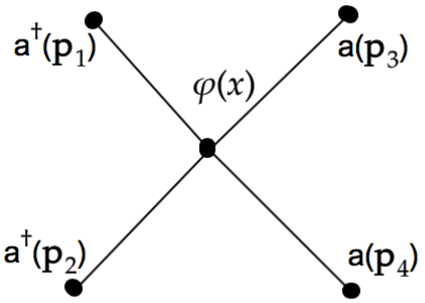

FIGURE 18.1. The diagrams for (18.2) are 4! repetitions of the above diagram.

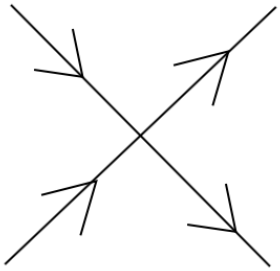

FIGURE 18.2. An example of a Feynman diagram, corresponding to the contraction diagram of Figure 18.1. Note that incoming particles enter from the left and exit on the right.

Now recall that for any $\mathbf{p} \in \mathbb{R}^3$, p denotes the vector $(\omega_{\mathbf{p}}, \mathbf{p}) \in X_m$. Any contraction diagram of the type shown in Figure 18.1 contributes

$$\begin{aligned}\langle 0|a(\mathbf{p}_3)\varphi(x)|0\rangle\langle 0|a(\mathbf{p}_4)\varphi(x)|0\rangle\langle 0|\varphi(x)a^\dagger(\mathbf{p}_1)|0\rangle\langle 0|\varphi(x)a^\dagger(\mathbf{p}_2)|0\rangle \\ = \frac{e^{i(x,p_3+p_4-p_1-p_2)}}{\sqrt{16\omega_{\mathbf{p}_1}\omega_{\mathbf{p}_2}\omega_{\mathbf{p}_3}\omega_{\mathbf{p}_4}}}.\end{aligned}$$

Multiplying the above by 4! and integrating over x, we get

$$\begin{aligned}\langle \mathbf{p}_3, \mathbf{p}_4|S|\mathbf{p}_1, \mathbf{p}_2\rangle &= -ig\int d^4x \frac{e^{i(x,p_3+p_4-p_1-p_2)}}{\sqrt{16\omega_{\mathbf{p}_1}\omega_{\mathbf{p}_2}\omega_{\mathbf{p}_3}\omega_{\mathbf{p}_4}}} + \mathcal{O}(g^2) \\ &= \frac{-ig(2\pi)^4\delta^{(4)}(p_3+p_4-p_1-p_2)}{\sqrt{16\omega_{\mathbf{p}_1}\omega_{\mathbf{p}_2}\omega_{\mathbf{p}_3}\omega_{\mathbf{p}_4}}} + \mathcal{O}(g^2). \qquad (18.3)\end{aligned}$$

What does (18.3) **mean?** Like in the Born approximation, we can conclude that (up to first order) the probability distribution of $(\mathbf{p}_3, \mathbf{p}_4)$, given that

the scattering has resulted in two outgoing particles, is supported on the manifold

$$\{(\mathbf{p}_3, \mathbf{p}_4) : p_3 + p_4 = p_1 + p_2\}$$

Note that this is a manifold in $\mathbb{R}^6$ described by 4 constraints. Therefore we expect this to be a 2D manifold. You can define a notion of 'Lebesgue measure' on this manifold as the limit of a sequence of measures with densities proportional to

$$\exp\left(-\frac{\|p_3 + p_4 - p_1 - p_2\|^2}{\varepsilon}\right)$$

as $\varepsilon \to 0$, where the multiplicative factor is taken in such a way as to give a nontrivial limit. The scattering amplitude implies that the conditional p.d.f. of $(\mathbf{p}_3, \mathbf{p}_4)$ with respect to this 'Lebesgue measure' on the manifold is proportional to

$$\frac{1}{\omega_{\mathbf{p}_3}\omega_{\mathbf{p}_4}}.$$

The constraint $p_3^0 + p_4^0 = p_1^0 + p_2^0$ shows that the manifold is bounded. Since the above density is also bounded, we can conclude that the density is integrable on the manifold and gives a legitimate probability measure.

The above reasoning can be made completely rigorous by replacing the improper incoming state $|\mathbf{p}_1, \mathbf{p}_2\rangle$ by a proper state which approximates it (for example, a Gaussian density), and then taking a sequence of approximations converging to the improper state.

Note that if we are in the non-relativistic limit, where $|\mathbf{p}_1| \ll m$ and $|\mathbf{p}_2| \ll m$, the constraint $p_3^0 + p_4^0 = p_1^0 + p_2^0$ approximately says

$$\frac{\mathbf{p}_3^2}{2m} + \frac{\mathbf{p}_4^2}{2m} = \frac{\mathbf{p}_1^2}{2m} + \frac{\mathbf{p}_2^2}{2m},$$

which is conservation of classical kinetic energy.

18.3. Words of caution

One should be aware of two things about the above calculation. Both have been mentioned before but are worth repeating here. First, the calculation is not rigorous because we do not know how to rigorously define φ^4 theory, or justify the Dyson expansion for this theory. However, if we ignore these two (severe) problems, the rest of the calculation can be easily made fully rigorous.

The second thing to be aware of is that φ^4 theory does not describe any known particle. It is purely a hypothetical theory that exhibits many of the complexities of quantum field theories that describe real particles, and is therefore useful for introducing various tools and techniques.

LECTURE 19

The Feynman propagator

Date: 11/5/2018
Scribe: Jimmy He

19.1. Towards a second order calculation in φ^4 theory

Continuing in the setting of the previous lecture, we will now calculate the second order term in the perturbative expansion for

$$\langle \mathbf{p}_3, \mathbf{p}_4 | S | \mathbf{p}_1, \mathbf{p}_2 \rangle \,,$$

that is, the term

$$\frac{(-ig)^2}{2!} \iint d^4x_1 d^4x_2 \, \langle 0 | a(\mathbf{p}_3) a(\mathbf{p}_4) \mathcal{T}(:\varphi(x_1)^4 :: \varphi(x_2)^4 :) a^\dagger(\mathbf{p}_1) a^\dagger(\mathbf{p}_2) | 0 \rangle \,.$$

From the theory we developed for Feynman diagrams, we know that we have to first put down vertices corresponding to each $\mathbf{p}_i$ and x_i, with 1 and 4 tails respectively, and we have to tie up all tails so that there are no untied tails. Moreover, Wick ordering stipulates that there cannot be self-loops at the x_i's. It is not hard to check that the only surviving Feynman diagrams are of the types shown in Figure 19.1.

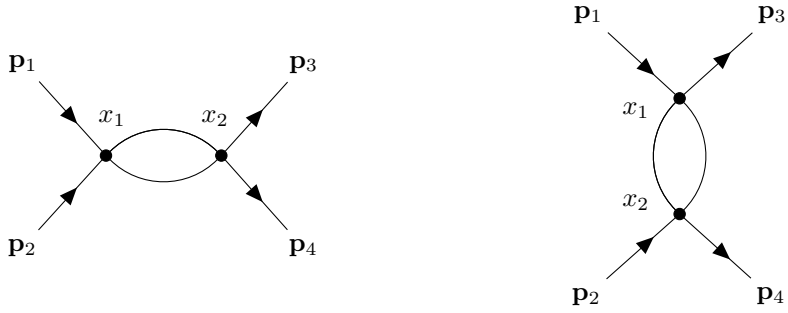

FIGURE 19.1. All Feynman diagrams appearing in the second order computation for φ^4 theory belong to one of the two types displayed above.

There are $2 \times (4 \times 3) \times (4 \times 3) \times 2$ diagrams of the first type shown in Figure 19.1, where the first 2 occurs because x_1 and x_2 can be switched, the 4×3 factors occur because we have to choose which tails of x_1 and x_2 should be tied to the incoming and outgoing momenta, and the last 2 occurs

because we have decide how to tie the remaining tails of x_1 and x_2 to each other. Similarly, there are $2 \times (4 \times 3) \times (4 \times 3) \times 2 \times 2$ diagrams of the second type, where the extra factor of 2 comes from being able to switch $\mathbf{p}_3$ and $\mathbf{p}_4$. These are called *one-loop diagrams* because there is only one loop.

Let us analyze the first diagram. This corresponds to the following product of contractions:

$$\begin{aligned}&\langle 0|\varphi(x_1)a^\dagger(\mathbf{p}_1)|0\rangle \langle 0|\varphi(x_1)a^\dagger(\mathbf{p}_2)|0\rangle \\ &\cdot (\langle 0|\mathcal{T}\varphi(x_1)\varphi(x_2))|0\rangle)^2 \langle 0|a(\mathbf{p}_3)\varphi(x_2))|0\rangle \langle 0|a(\mathbf{p}_4)\varphi(x_2))|0\rangle .\end{aligned}$$

We already know how to compute all but one of the terms in the above product, because we know that

$$\begin{aligned}\langle 0|a(\mathbf{p})\varphi(x)|0\rangle &= \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} e^{i(x,p)}, \\ \langle 0|\varphi(x)a^\dagger(\mathbf{p})|0\rangle &= \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} e^{-i(x,p)}.\end{aligned} \tag{19.1}$$

We will now see how to compute $\langle 0|\mathcal{T}\varphi(x_1)\varphi(x_2))|0\rangle$.

19.2. The Feynman propagator

Definition 19.1. The *Feynman propagator* is a tempered distribution on $\mathbb{R}^{1,3}$ formally defined as

$$\Delta_F(x) = i \int \frac{d^3\mathbf{p}}{(2\pi)^3 2\omega_{\mathbf{p}}} e^{-i|t|\omega_{\mathbf{p}} + i\mathbf{p}\cdot\mathbf{x}} \tag{19.2}$$

where $x = (t, \mathbf{x})$ and $\omega_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m^2}$.

Note that $\Delta_F(-x) = \Delta_F(x)$ because when we replace x by $-x$ in the integral, $|t|$ is unchanged and $\mathbf{p}\cdot\mathbf{x}$ becomes $-\mathbf{p}\cdot\mathbf{x}$, but we can do a change of variable $\mathbf{p} \mapsto -\mathbf{p}$ to retrieve the original form.

Proposition 19.1. *For any $x_1, x_2 \in \mathbb{R}^{1,3}$,*

$$\langle 0|\mathcal{T}\varphi(x_1)\varphi(x_2)|0\rangle = -i\Delta_F(x_1 - x_2).$$

Proof. Suppose that $x_1 = (t_1, \mathbf{x}_1)$ and $x_2 = (t_2, \mathbf{x}_2)$ with $t_1 \geq t_2$. Then

$$\langle 0|\mathcal{T}\varphi(x_1)\varphi(x_2)|0\rangle = \langle 0|\varphi(x_1)\varphi(x_2)|0\rangle .$$

Recall that $\varphi(x) = A(f_x) + A^\dagger(f_x)$ where $f_x = e^{i(x,p)}$. Thus by Lemma 17.1,

$$\begin{aligned}\langle 0|\varphi(x_1)\varphi(x_2)|0\rangle = (f_{x_1}, f_{x_2}) &= \int d\lambda_m(p) f_{x_1}(p)^* f_{x_2}(p) \\ &= \int d\lambda_m(p) e^{i(x_2 - x_1, p)} \\ &= \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{2\omega_{\mathbf{p}}} e^{i(x_2 - x_1, p)} = -i\Delta_F(x_1 - x_2).\end{aligned}$$

A similar derivation works if $t_1 < t_2$. □

19.3. An alternative expression for the Feynman propagator

The form of the Feynman propagator given in (19.2) is hard to work with, because of the presence of $\omega_{\mathbf{p}}$ in the denominator and in the exponent. Fortunately, it has a much friendlier form.

LEMMA 19.1. *As a tempered distribution,*

$$\Delta_F(x) = \lim_{\varepsilon \to 0^+} \int_{\mathbb{R}^{1,3}} \frac{d^4p}{(2\pi)^4} \frac{e^{-i(x,p)}}{-p^2 + m^2 - i\varepsilon}$$

PROOF. Let $x = (t, \mathbf{x})$ and let $p = (p^0, \mathbf{p})$. We first integrate the right side in p^0. Recall $p^2 = (p^0)^2 - \mathbf{p}^2$. So we want to compute:

$$\int_{-\infty}^{\infty} \frac{dp^0}{2\pi} \frac{e^{-itp^0}}{-(p^0)^2 + \mathbf{p}^2 + m^2 - i\varepsilon}$$

Let's write $z = p^0$ and $f(z) = -z^2 + \mathbf{p}^2 + m^2 - i\varepsilon$, so that we have to compute

$$\int_{-\infty}^{\infty} \frac{dz}{2\pi} \frac{e^{-itz}}{f(z)}.$$

We will calculate this integral using contour integration. For that, it is important to understand the behavior of the quadratic polynomial f near its roots. If $\varepsilon = 0$, the roots of f are $\pm\omega_{\mathbf{p}}$. For $\varepsilon > 0$, the roots are $\pm\omega_{\mathbf{p},\varepsilon}$, where $\omega_{\mathbf{p},\varepsilon}$ is slightly below $\omega_{\mathbf{p}}$ in the complex plane, and $-\omega_{\mathbf{p},\varepsilon}$ is slightly above $-\omega_{\mathbf{p}}$.

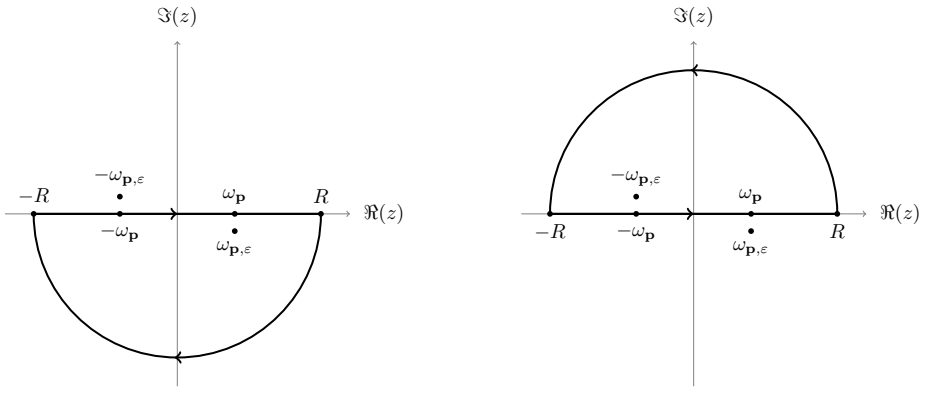

FIGURE 19.2. Contours for $t \geq 0$ and $t < 0$.

Suppose that $t \geq 0$. Then we take a contour going from $-R$ to R along the real axis and then back to $-R$ along a semicircle below the real axis (the left side of Figure 19.3). Since $t \geq 0$, we can show that the contribution of the semicircular part approaches 0 as $R \to \infty$. If $t < 0$, we take the flipped contour going above the real axis (the right side of Figure 19.3). There is

only one pole that we have to consider for each case, and the residues for the two poles are

$$-\frac{e^{-it\omega_{\mathbf{p},\varepsilon}}}{2\omega_{\mathbf{p},\varepsilon}} \quad \text{at } \omega_{\mathbf{p},\varepsilon}, \qquad\qquad \frac{e^{it\omega_{\mathbf{p},\varepsilon}}}{2\omega_{\mathbf{p},\varepsilon}} \quad \text{at } -\omega_{\mathbf{p},\varepsilon}.$$

So using Cauchy's theorem gives

$$\int_{-\infty}^{\infty} \frac{dz}{2\pi} \frac{e^{-itz}}{f(z)} = \begin{cases} i(2\omega_{\mathbf{p},\varepsilon})^{-1} e^{-it\omega_{\mathbf{p},\varepsilon}} & \text{if } t \geq 0, \\ i(2\omega_{\mathbf{p},\varepsilon})^{-1} e^{it\omega_{\mathbf{p},\varepsilon}} & \text{if } t < 0. \end{cases}$$

This completes the proof. □

19.4. Putting it all together

Using Lemma 19.1, we get

$$\begin{aligned} &\langle 0|\mathcal{T}\varphi(x_1)\varphi(x_2)|0\rangle^2 = (-i\Delta_F(x_1 - x_2))^2 \\ &= -\lim_{\varepsilon\to 0^+} \iint \frac{d^4p d^4p'}{(2\pi)^8} \frac{e^{-i(x_1-x_2,p+p')}}{(-p^2+m^2-i\varepsilon)(-p'^2+m^2-i\varepsilon)}. \end{aligned}$$

Putting this together with (19.1), we obtain

$$\begin{aligned} &\langle 0|\varphi(x_1)a^\dagger(\mathbf{p}_1)|0\rangle \langle 0|\varphi(x_1)a^\dagger(\mathbf{p}_2)|0\rangle \\ &\cdot (\langle 0|\mathcal{T}\varphi(x_1)\varphi(x_2))|0\rangle)^2 \langle 0|a(\mathbf{p}_3)\varphi(x_2))|0\rangle \langle 0|a(\mathbf{p}_4)\varphi(x_2))|0\rangle \\ &= \frac{-1}{\sqrt{16\omega_{\mathbf{p}_1}\omega_{\mathbf{p}_2}\omega_{\mathbf{p}_3}\omega_{\mathbf{p}_4}}} \\ &\qquad \cdot \lim_{\varepsilon\to 0^+} \iint \frac{d^4p d^4p'}{(2\pi)^8} \frac{e^{i(x_2,p_3+p_4+p+p')} e^{-i(x_1,p_1+p_2+p+p')}}{(-p^2+m^2-i\varepsilon)(-p'^2+m^2-i\varepsilon)}. \end{aligned}$$

We will continue from here in the next lecture.

LECTURE 20

The problem of infinities

Date: 11/7/2018
Scribe: Sohom Bhattacharya

20.1. Completing the second order calculation in φ^4 theory

Let us continue from where we stopped in the previous lecture. Recall that we were trying to calculate the second order term in the perturbative expansion for a scattering amplitude in φ^4 theory, and we ended up with a term containing the integral

$$\lim_{\varepsilon\to 0^+}\int\int \frac{d^4pd^4p'}{(2\pi)^8}\frac{e^{i(x_2,p_3+p_4+p+p')}e^{-i(x_1,p_1+p_2+p+p')}}{(-p^2+m^2-i\varepsilon)(-(p')^2+m^2-i\varepsilon)}.$$

To get the second order term in the Dyson series, we have to integrate this with respect to x_1 and x_2. Note that

$$\begin{aligned}&\iint d^4x_1d^4x_2e^{i(x_2,p_3+p_4+p+p')-i(x_1,p_1+p_2+p+p')}\\&=(2\pi)^8\delta^{(4)}(p_3+p_4+p+p')\delta^{(4)}(p_1+p_2+p+p').\end{aligned}$$

Recall the identity

$$\int_{-\infty}^{\infty}\delta(x-z)\delta(y-z)\xi(z)dz=\delta(x-y)\xi(x),$$

which holds true in higher dimensions also. Using this identity, and exchanging integrals and limits at will, we get

$$\begin{aligned}&\lim_{\varepsilon\to 0+}\iiiint d^4x_1d^4x_2\int\frac{d^4p}{(2\pi)^4}\frac{d^4p'}{(2\pi)^4}\frac{e^{i(x_2,p_3+p_4+p+p')-i(x_1,p_1+p_2+p+p')}}{(-p^2+m^2-i\varepsilon)(-p'^2+m^2-i\varepsilon)}\\&=\lim_{\varepsilon\to 0+}\iint d^4pd^4p'\frac{\delta^{(4)}(p_3+p_4+p+p')\delta^{(4)}(p_1+p_2+p+p')}{(-p^2+m^2-i\varepsilon)(-p'^2+m^2-i\varepsilon)}\\&=\lim_{\varepsilon\to 0+}\int d^4p\frac{\delta^{(4)}(p_3+p_4-p_1-p_2)}{(-p^2+m^2-i\varepsilon)(-(-p_1-p_2-p)^2+m^2-i\varepsilon)},\end{aligned}$$

which finally by a change of variable $p\mapsto -p$ yields

$$\lim_{\varepsilon\to 0+}\int d^4p\frac{\delta^{(4)}(p_3+p_4-p_1-p_2)}{(-p^2+m^2-i\varepsilon)(-(p_1+p_2-p)^2+m^2-i\varepsilon)}.$$

Now let

$$M_1 = \lim_{\varepsilon \to 0+} \int \frac{d^4p}{(2\pi)^4} \frac{1}{(-p^2 + m^2 - i\varepsilon)(-(p_1 + p_2 - p)^2 + m^2 - i\varepsilon)},$$

which is just the previous expression but without the Dirac delta, and divided by $(2\pi)^4$. Recalling the discussion from the previous lecture and the first-order computation we did before, we now see that

$$\begin{aligned} &\langle \mathbf{p}_3, \mathbf{p}_4 | S | \mathbf{p}_1, \mathbf{p}_2 \rangle \\ &= (g - ig^2 M + \mathcal{O}(g^3)) \frac{-i(2\pi)^4 \delta^{(4)}(p_3 + p_4 - p_1 - p_2)}{\sqrt{16 \omega_{\mathbf{p_1}} \omega_{\mathbf{p_2}} \omega_{\mathbf{p_3}} \omega_{\mathbf{p_4}}}}, \end{aligned} \tag{20.1}$$

where M is the sum of several terms like M_1, multiplied by some appropriate combinatorial factor. (It is not hard to argue that the $\delta^{(4)}$ and the ω's would occur in every term of the perturbative expansion just as they did in the first and second terms, which allows us to bring the $\mathcal{O}(g^3)$ term inside the bracket.)

The big issue with the integral defining M_1 is that the integrand decays like $\|p\|^{-4}$ as $\|p\| \to \infty$ and this makes the integral divergent. This is the simplest instance of the 'problem of infinities' in quantum field theory.

20.2. The idea of renormalization

The quantity we are interested in is $g - ig^2 M$. The theory gives M as a divergent integral. However, experiment gives M as a finite quantity (strictly speaking, not in φ^4 theory, but in models that represent real phenomena). Therefore, obviously there is something wrong with the theory. The optimistic viewpoint is that the theory is approximately correct, which means that the integral which diverged must be replaced with an integral which converges by multiplying the integrand by a cutoff function, which is given by nature but unknown to us. For example, in our context, this means that M should be replaced by something like

$$\begin{aligned} \widetilde{M} = \lim_{\varepsilon \to 0+} \int \frac{d^4p}{(2\pi)^4} &\frac{\theta(\mathbf{p})}{(-p^2 + m^2 - i\varepsilon)(-(p_1 + p_2 - p)^2 + m^2 - i\varepsilon)} \\ &+ \text{ other similar terms,} \end{aligned}$$

where $\theta(\mathbf{p})$ is a function that always takes values between 0 and 1, satisfies $\theta(\mathbf{p}) = 1$ when $|\mathbf{p}| \leq R$ for some large R, and decays fast enough in the region $|\mathbf{p}| > R$ so that the integral converges. Further, we assume that g is so small that we can effectively treat $\widetilde{M}$ as a large but fixed value while working out the perturbative expansion in g. Thus, terms like $\mathcal{O}(g^k)$ may implicitly contain powers of quantities like $\widetilde{M}$.

The problem is that we do not know the cutoff function θ that is provided by nature. This hurdle is overcome as follows, which is the main idea in renormalization. Suppose that we can observe the value of $\widetilde{M}$ for certain values $\mathbf{p}_1^*$, $\mathbf{p}_2^*$, $\mathbf{p}_3^*$, $\mathbf{p}_4^*$ of the incoming and outgoing momenta. For example,

they may all be close to $(m, 0, 0, 0)$, which is the non-relativistic scenario realized in a laboratory experiment. Call this value $\widetilde{M}^*$. Then the theory will allow you to calculate $\widetilde{M}$ for any $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$, $\mathbf{p}_4$, *even if you do not know the cutoff function or the coupling constant*, through a process called renormalization. This is roughly described as follows, for the problem that we are discussing.

For the laboratory values $\mathbf{p}_1^*$, $\mathbf{p}_2^*$, $\mathbf{p}_3^*$, $\mathbf{p}_4^*$, suppose that the true observed amplitude is

$$A_* \frac{-i(2\pi)^4\delta^{(4)}(p_3^* + p_4^* - p_1^* - p_2^*)}{\sqrt{16\omega_{\mathbf{p}_1^*}\omega_{\mathbf{p}_2^*}\omega_{\mathbf{p}_3^*}\omega_{\mathbf{p}_4^*}}}.$$

(Here we assume that $p_3^* + p_4^* - p_1^* - p_2^* = 0$, so that A_* can be measured.) Then by (20.1),

$$A_* = g - ig^2\widetilde{M}^* + \mathcal{O}(g^3), \tag{20.2}$$

where $\widetilde{M}^*$ is the value of $\widetilde{M}$ in the laboratory setting. Also, for our given $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$, $\mathbf{p}_4$,

$$\langle \mathbf{p}_3, \mathbf{p}_4 | S | \mathbf{p}_1, \mathbf{p}_2 \rangle = A \frac{-i(2\pi)^4\delta^{(4)}(p_3 + p_4 - p_1 - p_2)}{\sqrt{16\omega_{\mathbf{p_1}}\omega_{\mathbf{p_2}}\omega_{\mathbf{p_3}}\omega_{\mathbf{p_4}}}},$$

where

$$A = g - ig^2\widetilde{M} + \mathcal{O}(g^3). \tag{20.3}$$

The quantity A is our main object of interest, which we cannot evaluate if we do not know the cutoff function θ. We will now argue that in fact we can approximately evaluate A even if do not know θ, purely using the amplitude A_* observed in the laboratory. To see this, first note that by (20.2),

$$g = A_* + ig^2\widetilde{M}^* + \mathcal{O}(g^3). \tag{20.4}$$

Next, note that since $A_* = \mathcal{O}(g)$, squaring the equation (20.2) we obtain

$$g^2 = A_*^2 + \mathcal{O}(g^3). \tag{20.5}$$

Substituting the values of g and g^2 from (20.4) and (20.5) into (20.3), we get

$$\begin{aligned} A &= A_* + ig^2\widetilde{M}^* - ig^2\widetilde{M} + \mathcal{O}(g^3) \\ &= A_* - iA_*^2(\widetilde{M} - \widetilde{M}^*) + \mathcal{O}(g^3). \end{aligned} \tag{20.6}$$

It turns out that if the cutoff function θ satisfies $\theta(\mathbf{p}) \in [0, 1]$ for all $\mathbf{p}$, $\theta(\mathbf{p}) = 1$ for $|\mathbf{p}| \leq R$, and decays fast enough the region $|\mathbf{p}| > R$, then the limit

$$L := \lim_{R\to\infty} (\widetilde{M} - \widetilde{M}^*)$$

exists and is finite, irrespective of how θ varies with R, as long as it satisfies the above constraints. (We will prove this in the next couple of lectures.) We can then substitute L for $\widetilde{M} - \widetilde{M}^*$ in (20.6) to get an approximate value of A in terms of A_*, with $\mathcal{O}(g^3)$ error. This is the key idea in perturbative

renormalization, which becomes harder to manage for higher orders of perturbation theory. If, however, it can be done for all orders, the theory is called 'perturbatively renormalizable'.

Note that the above calculation was slightly complicated because we wanted to get an approximation for A solely in terms of A_*, and not using any information about θ or g. If we know g, then the problem becomes easier, because we can simply approximate $\widetilde{M}$ by $L + \widetilde{M}^*$ in (20.3).

LECTURE 21

One-loop renormalization in φ^4 theory

Date: 11/9/2018
Scribe: Laura Lyman

21.1. A toy example

Recall the basic idea of renormalization from the previous lecture: Suppose that we want to carry out a calculation for a physical system, where we input some quantity a, where a can be a vector or scalar (e.g. the 4-tuple of incoming and outgoing momenta $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$, $\mathbf{p}_4$, as considered in the previous lecture), and the output is a scalar $f(a)$ (e.g. the probability amplitude $\langle \mathbf{p}_3, \mathbf{p}_4 | S | \mathbf{p}_1, \mathbf{p}_2 \rangle$). As seen in the previous lecture, sometimes the theory will yield a prediction for $f(a)$ in terms of divergent integrals. However the observed value of $f(a)$ is finite. The optimistic viewpoint is that the theory is approximately correct, in the sense that the divergent integrals should be replaced with integrals with some cutoffs (*regularized* versions). Nature provides the cutoff function θ, but it is unknown to us. The solution to this obstacle is that we can still approximately recover $f(a)$ if we know $f(a')$ for any single a' *even if* θ *is unknown.* When this can be done for all orders of perturbation theory, the theory is called *perturbatively renormalizable.*

To understand the situation, consider the following toy example. Suppose that the input quantity is some number $a > 0$, and the output predicted by theory is

$$f_{theory}(a) = \int_0^\infty \frac{dx}{x+a} = \infty.$$

However, suppose that the experimentally observed output $f(a)$ is always finite. To resolve this discrepancy, assume that the theory is approximately correct, in the sense that the true $f(a)$ is given by

$$f(a) = \int_0^\infty \frac{\theta(x)\,dx}{x+a}$$

where θ is a function such that $\theta(x) \in [0,1]$ for all x, $\theta(x) = 1$ when $x \leq R$ for some large $R \in \mathbb{R}$, and θ decays sufficiently fast in the region $x > R$ so that the integral converges. Then

$$f(a) - f(a') = \int_0^\infty \frac{\theta(x)(a'-a)\,dx}{(x+a)(x+a')} = \int_0^\infty \frac{dx(a-a')}{(x+a)(x+a')} + \mathcal{O}(1/R).$$

Thus, we can approximately recover $f(a)$ for any a if we observe the value of $f(a')$ for some a', even if we do not know θ.

21.2. Main result

Recall that we were trying to calculate the second order term in the Dyson expansion of

$$\langle \mathbf{p_3}, \mathbf{p_4} | S | \mathbf{p_1}, \mathbf{p_2} \rangle.$$

There were no divergent integrals in the first order term. However, in the second order, each Feynman diagram had a single loop, which gave divergent integrals like

$$\lim_{\varepsilon \to 0^+} \int \frac{d^4p}{(2\pi)^4} \frac{1}{(-p^2 + m^2 - i\varepsilon)(-(w-p)^2 + m^2 - i\varepsilon)}$$

where $w = p_1 + p_2$. Let us therefore assume that the "true" value is

$$U(w, \theta) := \lim_{\varepsilon \to 0^+} \int \frac{d^4p}{(2\pi)^4} \frac{\theta(\mathbf{p})}{(-p^2 + m^2 - i\varepsilon)(-(w-p)^2 + m^2 - i\varepsilon)}$$

where $\theta : \mathbb{R}^3 \to [0,1]$ is a function such that $\theta(\mathbf{p}) = 1$ when $|\mathbf{p}| \leq R$, and θ decays sufficiently fast in the region $|\mathbf{p}| > R$. The following theorem shows that the above integral can be approximately computed for any w if we know its value at a single w', even if we do not know θ.

THEOREM 21.1 (One-loop renormalization). *For any $w, w' \in \mathbb{R}^{1,3}$*

$$\lim_{R \to \infty} (U(w, \theta) - U(w', \theta))$$

exists, is finite, and depends only on (w, w'). Here we assume that θ varies with R in such a way that we always have $\theta(\mathbf{p}) = 1$ for $|\mathbf{p}| \leq R$.

Why is the situation of Theorem 21.1 harder to analyze than the toy example discussed above? The difference is that terms like $-p^2 + m^2$ in the denominator introduce infinite manifolds of singularities as $\varepsilon \to 0$. To get rid of such singularities, we need two technical tools.

LEMMA 21.1 (Feynman parameter). *Suppose that $A, B \in \mathbb{C}$ are such that the line segment joining A and B in the complex plane does not pass through 0. Then*

$$\frac{1}{AB} = \int_0^1 du \frac{1}{(Au + B(1-u))^2}.$$

(Here u is called a Feynman parameter.)

PROOF. Note that

$$\frac{d}{du}\left[\frac{1}{(B-A)(Au + B(1-u))}\right] = \frac{1}{(Au + B(1-u))^2},$$

and substitute above. □

LEMMA 21.2 (Wick rotation). *Suppose that $f : \mathbb{C} \to \mathbb{C}$ is a holomorphic function which has no poles in the region $\{x + iy : xy \geq 0\}$. Furthermore, suppose that $|f(z)| \leq C|z|^{-1-\varepsilon}$ in this region for some $C < \infty$ and $\varepsilon > 0$. Then*

$$\int_{-\infty}^{\infty} dx f(x) = i \int_{-\infty}^{\infty} dx f(ix).$$

PROOF. To see this, consider the following contour. Starting from the origin, traverse the blue path and then the red path in Figure 21.1.

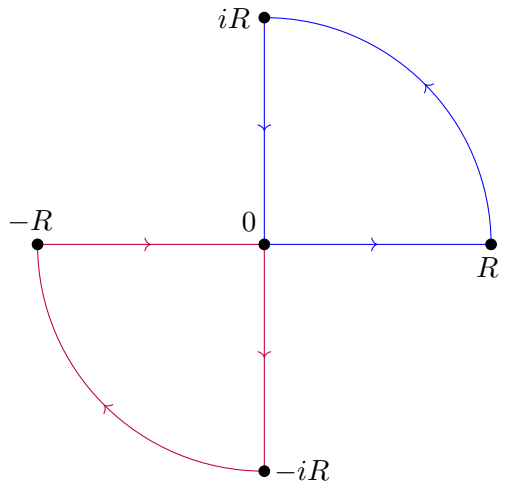

FIGURE 21.1. Contour traversed from the origin first along the blue path and then along the red path.

Note that f has no poles enclosed by this contour. The decay condition implies that the integral contribution on the circular arcs tends to 0 as $R \to \infty$. Thus, the result follows by Cauchy's theorem and taking $R \to \infty$. □

COROLLARY 21.1. *For any $\alpha > 0$, $E \in \mathbb{R}$, and $\varepsilon > 0$,*

$$\int_{-\infty}^{\infty} dx \frac{1}{(-\alpha x^2 + E - i\varepsilon)^2} = i \int_{-\infty}^{\infty} dx \frac{1}{(\alpha x^2 + E - i\varepsilon)^2}.$$

PROOF. Since $\varepsilon > 0$, it is easy to see that the integrals are both well-defined and finite. Let

$$f(z) = \frac{1}{(-\alpha z^2 + E - i\varepsilon)^2}.$$

Then there is some constant C such that $|f(z)| \leq C|z|^{-2}$ for all $|z|$. Moreover, observe that

$$-\alpha z^2 = -\alpha(x^2 - y^2 + 2ixy) = -\alpha(x^2 - y^2) - 2i\alpha xy.$$

So if $xy \geq 0$, then $\Im(-\alpha z^2) \leq 0$, and therefore $-\alpha z^2 + E - i\varepsilon \neq 0$. Thus we can apply Wick rotation (Lemma 21.2) and obtain the desired result. $\square$

21.3. Towards the proof of the main result

Let us now apply the above techniques to $U(w, \theta)$ to start the proof of Theorem 21.1. First observe that by the Feynman parameter trick,

$$\begin{aligned} U(w,\theta) &= \lim_{\varepsilon \to 0^+} \int \frac{d^4p}{(2\pi)^4} \frac{\theta(\mathbf{p})}{-(p^2 + m^2 - i\varepsilon)(-(w-p)^2 + m^2 - i\varepsilon)} \\ &= \lim_{\varepsilon \to 0^+} \int_0^1 du \int \frac{d^4p}{(2\pi)^4} \frac{\theta(\mathbf{p})}{(-(w-p)^2 u - p^2(1-u) + m^2 - i\varepsilon)^2}. \end{aligned}$$

By the change of variable $p \to p + uw$, this equals

$$= \lim_{\varepsilon \to 0^+} \int_0^1 du \int \frac{d^4p}{(2\pi)^4} \frac{\theta(\mathbf{p} + u\mathbf{w})}{(-p^2 - u(1-u)w^2 + m^2 - i\varepsilon)^2}.$$

Let us first integrate in p^0, fixing u and $\mathbf{p}$. By Wick rotation (Corollary 21.1), we get

$$\begin{aligned} &\int_{-\infty}^{\infty} \frac{dp^0}{2\pi} \frac{1}{(-(p^0)^2 + p^2 - u(1-u)w^2 + m^2 - i\varepsilon)^2} \\ &= i \int_{-\infty}^{\infty} \frac{dp^0}{2\pi} \frac{1}{(\|p\|^2 - u(1-u)w^2 + m^2 - i\varepsilon)^2}, \end{aligned}$$

where $\|p\|$ is the Euclidean norm of p. Thus,

$$U(w,\theta) = i \lim_{\varepsilon \to 0^+} \int_0^1 du \int \frac{d^4p}{(2\pi)^4} \frac{\theta(\mathbf{p} + u\mathbf{w})}{(\|p\|^2 - u(1-u)w^2 + m^2 - i\varepsilon)^2}.$$

Let us now write

$$U(w,\theta) = U_1(w,\theta) + U_2(w,\theta)$$

where U_1 is the integral over $\|p\| \leq R/2$ and U_2 is the rest of the integral.

LEMMA 21.3. *If R is sufficiently large, we have*

$$U_1(w,\theta) = \frac{i}{16\pi^2} \int_0^1 du A(u),$$

where

$$\begin{aligned} A(u) = \log((R/2)^2 - u(1-u)w^2 + m^2) - \log(-u(1-u)w^2 + m^2) \\ - \frac{(R/2)^2}{(R/2)^2 - u(1-u)w^2 + m^2}, \end{aligned}$$

where we use the convention $\log(-x) = \log x - i\pi$ *for* $x > 0$.

PROOF. If R is sufficiently large, note that $\theta(\mathbf{p}+u\mathbf{w}) = 1$ for any p such that $\|p\| \leq R/2$, and any $u \in [0,1]$. Thus, letting $C = -u(1-u)w^2 + m^2 - i\varepsilon$ and using polar coordinates, we get

$$\begin{aligned} U_1(w,\theta) &= i \lim_{\varepsilon \to 0^+} \int_0^1 du \int_{\|p\| \leq R/2} \frac{d^4p}{(2\pi)^4} \frac{1}{(\|p\|^2 + C)^2} \\ &= \lim_{\varepsilon \to 0^+} \frac{i}{16\pi^4} \int_0^1 du \ (2\pi^2) \int_0^{R/2} \frac{r^3 dr}{(r^2+C)^2}. \end{aligned}$$

Now

$$\begin{aligned} \int_0^{R/2} \frac{r^3 dr}{(r^2+C)^2} &= \int_0^{R/2} \frac{r dr}{r^2+C} - C \int_0^{R/2} \frac{r dr}{(r^2+C)^2} \\ &= \frac{1}{2}\left[\log((R/2)^2 + C) - \log C\right] + \frac{C}{2}\left[\frac{1}{(R/2)^2+C} - \frac{1}{C}\right] \\ &= \frac{1}{2}\left[\log((R/2)^2 + C) - \log C - \frac{(R/2)^2}{(R/2)^2+C}\right], \end{aligned}$$

where we use the branch of the logarithm that is defined on $\mathbb{C} \setminus (-\infty, 0]$. Now sending $\varepsilon \to 0^+$ in the definition of C completes the proof. $\square$

LECTURE 22

A glimpse at two-loop renormalization

Date: 11/12/2018
Scribe: Lingxiao Li

22.1. Finishing the proof for one-loop renormalization

In this lecture we will first complete the proof of Theorem 21.1. Recall the quantity $U(w,\theta)$ from the previous lecture and the decomposition $U = U_1 + U_2$. From the expression we obtained for $U_1(w,\theta)$, it is easy to see that

$$\lim_{R\to\infty}(U_1(w,\theta) - U_1(w',\theta)) = L(w) - L(w'),$$

where

$$L(w) = \frac{-i}{16\pi^2}\int_0^1 du \log(m^2 - u(1-u)w^2).$$

As before, here we use the convention that $\log(-x) = \log x - i\pi$ when $x > 0$. Our next step is to prove the following lemma.

Lemma 22.1.

$$\lim_{R\to\infty}(U_2(w,\theta) - U_2(w',\theta)) = 0.$$

This will imply Theorem 21.1, and will moreover prove that

$$\lim_{R\to\infty}(U(w,\theta) - U(w',\theta)) = L(w) - L(w').$$

Proof of Lemma 22.1. By definition, we have

$$U_2(w,\theta) = i\lim_{\varepsilon\to 0^+}\int_0^1 du \int_{\|p\|\geq R/2}\frac{d^4p}{(2\pi)^4}\frac{\theta(\mathbf{p}+u\mathbf{w})}{(\|p\|^2 - u(1-u)w^2 + m^2 - i\varepsilon)^2}.$$

If $\|p\|$ is sufficiently large (by letting R be large), the denominator in the integrand will be far away from 0, so we can interchange limit and integrals and send $\varepsilon \to 0$ to get

$$\begin{aligned} U_2(w,\theta) &= i\int_0^1 du \int_{\|p\|\geq R/2}\frac{d^4p}{(2\pi)^4}\frac{\theta(\mathbf{p}+u\mathbf{w})}{(\|p\|^2 - u(1-u)w^2 + m^2)^2} \\ &= i\int_0^1 du \int_{\|p-uw\|\geq R/2}\frac{d^4p}{(2\pi)^4}\frac{\theta(\mathbf{p})}{(\|p-uw\|^2 - u(1-u)w^2 + m^2)^2}, \end{aligned}$$

where the last step follows by the change of variable $p \mapsto p - uw$. Let us further split the integral as

$$U_3(w,\theta) = i\int_0^1 du \int_{\|p\|\geq R/2} \frac{d^4p}{(2\pi)^4} \frac{\theta(\mathbf{p})}{(\|p-uw\|^2 - u(1-u)w^2 + m^2)^2},$$

and $U_4(w,\theta) = U_2(w,\theta) - U_3(w,\theta)$. Let $\Delta(w,R)$ be the symmetric difference of the regions $\{p : \|p\| \geq R/2\}$ and $\{p : \|p-uw\| \geq R/2\}$. Since $\theta \in [0,1]$ everywhere, it follows that

$$|U_4(w,\theta)| \leq \int_0^1 du \int_{\Delta(w,R)} \frac{d^4p}{(2\pi)^4} \frac{1}{(\|p-uw\|^2 - u(1-u)w^2 + m^2)^2}.$$

Since $\mathrm{Vol}(\Delta(w,R)) = \mathcal{O}(R^3)$ and the integrand in the above display is of order R^{-4} in $\Delta(w,R)$, it follows that $|U_4(w,\theta)| = \mathcal{O}(R^{-1})$ as $R \to \infty$. So to complete the proof of the lemma, it remains to show that

$$\lim_{R\to\infty} (U_4(w,\theta) - U_4(w',\theta)) = 0.$$

To prove this, note that

$$\begin{aligned} |U_4(w,\theta) - U_4(w',\theta)| &\leq \int_0^1 du \int_{\|p\|\geq R/2} \frac{d^4p}{(2\pi)^4} |\theta(\mathbf{p}) F(p,u,w,w')| \\ &\leq \int_0^1 du \int_{\|p\|\geq R/2} \frac{d^4p}{(2\pi)^4} |F(p,u,w,w')|, \end{aligned}$$

where

$$\begin{aligned} F(p,u,w,w') = &\frac{1}{(\|p\|^2 - u(1-u)w^2 + m^2)^2} \\ &- \frac{1}{(\|p\|^2 - u(1-u)w'^2 + m^2)^2}. \end{aligned}$$

It is easy to see that there is some R_0 and some C, depending only on m, w and w', such that $|F(p,u,w,w')| \leq C\|p\|^{-5}$ whenever $\|p\| \geq R_0$. Thus,

$$\lim_{R\to\infty} |U_4(w,\theta) - U_4(w',\theta)| \leq \lim_{R\to\infty} \int_0^1 du \int_{\|p\|\geq R/2} \frac{d^4p}{(2\pi)^4} C\|p\|^{-5} = 0.$$

This completes the proof of the lemma and hence of Theorem 21.1. □

22.2. Wick ordering as a form of renormalization

Suppose we use $\varphi(x)^4$ in the Hamiltonian of φ^4 theory instead of $:\varphi(x)^4:$. Then we will end up getting many more exploding integrals, like the ones here, even in the first order of perturbation theory. But after renormalization, all these infinities will cancel out, and we will get exactly the same predictions as we did in the Wick ordered version. Hence Wick ordering is like a 'preemptive' renormalization.

22.3. The counterterm method

The counterterm method is a different (and more conventional) approach for doing the same calculations that we did above. The counterterm method involves the following:

- Assume a form of the cutoff.
- Reparametrize the model parameters (in our case, only g) depending on the cutoff and express the new Hamiltonian as the old Hamiltonian plus some extra terms involving the new parameters, which are called counterterms.
- The reparametrization should be done in such a way that the results of the calculations automatically appear in renormalized form.

To clarify the above ideas, let us consider the example of one-loop renormalization in φ^4 theory. Fix a cutoff and recall the quantities $\widetilde{M}$ and $\widetilde{M}^*$ that appeared in our calculations in Section 20.2. Let us reparametrize our 'bare' coupling constant g as

$$g = g_r + ig_r^2\widetilde{M}^*,$$

where g_r is a new parameter, and the subscript r stands for 'renormalized'. Then notice that our quantity of interest from Section 20.2 is

$$\begin{aligned} g - ig^2\widetilde{M} + \mathcal{O}(g^3) &= (g_r + ig_r^2\widetilde{M}^*) - i(g_r + ig_r^2\widetilde{M}^*)^2\widetilde{M} \\ &= g_r - ig_r^2(\widetilde{M} - \widetilde{M}^*) + \mathcal{O}(g_r^3). \end{aligned}$$

By Theorem 21.1, we now know that $\widetilde{M} - \widetilde{M}^*$ approaches a finite limit as the cutoff is removed. Thus, when the bare coupling parameter g is reparametrized in terms of the renormalized parameter g_r, there are no infinities in the second order of perturbation theory. This reparametrization is conventionally described by saying that we add the counterterm $\frac{1}{4!}ig_r^2\widetilde{M}^*{:}\varphi(x)^4{:}$ to the interaction Hamiltonian density $\frac{1}{4!}g_r{:}\varphi(x)^4{:}$. Note that the counterterm depends on the choice of the cutoff, as well the choice of the momenta $\mathbf{p}_1^*$, $\mathbf{p}_2^*$, $\mathbf{p}_3^*$, $\mathbf{p}_4^*$ that we chose for our hypothetical laboratory experiment in Section 20.2.

22.4. A glimpse at two-loop renormalization

For the third order of perturbation in φ^4 theory, we will end up with an expression like

$$\begin{aligned} &\langle \mathbf{p}_3, \mathbf{p}_4 | S | \mathbf{p}_1, \mathbf{p}_2 \rangle \\ &= (g - ig^2 M_2 + g^3 M_3 + \mathcal{O}(g^4)) \frac{-i(2\pi)^4 \delta^{(4)}(p_3 + p_4 - p_1 - p_2)}{\sqrt{16\omega_{\mathbf{p}_1}\omega_{\mathbf{p}_1}\omega_{\mathbf{p}_3}\omega_{\mathbf{p}_4}}}, \end{aligned}$$

where M_2 and M_3 are expressed as sums of divergent integrals. As in Section 20.2, we assume that these are only approximations, and the true values

are $\widetilde{M}_2$ and $\widetilde{M}_3$, obtained by cutting off the divergent integrals in some suitable way. We are interested in evaluating the amplitude

$$A = g - ig^2\widetilde{M}_2 + g^3\widetilde{M}_3 + \mathcal{O}(g^4). \tag{22.1}$$

Suppose that A_* is the laboratory value of this amplitude, for incoming and outgoing momenta $\mathbf{p}_1^*$, $\mathbf{p}_2^*$, $\mathbf{p}_3^*$, $\mathbf{p}_4^*$. Express A_* as

$$A_* = g - ig^2\widetilde{M}_2^* + g^3\widetilde{M}_3^* + \mathcal{O}(g^4). \tag{22.2}$$

Then how can we recover A from A_*? Recall that in one-loop renormalization, we got

$$A = A_* - iA_*^2(\widetilde{M}_2 - \widetilde{M}_2^*) + \mathcal{O}(g^3),$$

and then we showed that $\widetilde{M}_2 - \widetilde{M}_2^*$ has a finite limit as the cutoff is removed. The story for two-loop renormalization is quite a bit more complicated. First, note that by (22.2),

$$g = A_* + ig^2\widetilde{M}_2^* - g^3\widetilde{M}_3^* + \mathcal{O}(g^4) \tag{22.3}$$

Squaring both sides, and using the fact that $A_* = \mathcal{O}(g)$, we get

$$g^2 = A_*^2 + 2ig^2A_*\widetilde{M}_2^* + \mathcal{O}(g^4).$$

The above expression shows, in particular, that $g^2 = A_*^2 + \mathcal{O}(g^3)$. Plugging this in place of the g^2 on the right and using $A_* = \mathcal{O}(g)$ gives

$$g^2 = A_*^2 + 2iA_*^3\widetilde{M}_2^* + \mathcal{O}(g^4). \tag{22.4}$$

Cubing both sides of (22.3) and again using $A_* = \mathcal{O}(g)$, we have

$$g^3 = A_*^3 + \mathcal{O}(g^4). \tag{22.5}$$

Using the values of g^2 and g^3 from (22.4) and (22.5) in (22.3) gives

$$g = A_* + iA_*^2\widetilde{M}_2^* - A_*^3(2(\widetilde{M}_2^*)^2 + \widetilde{M}_3^*) + \mathcal{O}(g^4). \tag{22.6}$$

Plugging in the values of g, g^2 and g^3 obtained in (22.6), (22.4) and (22.5) into equation (22.1), we get

$$\begin{aligned} A &= A_* + iA_*^2\widetilde{M}_2^* - A_*^3(2(\widetilde{M}_2^*)^2 + \widetilde{M}_3^*) \\ &\quad - i(A_*^2 + 2iA_*^3\widetilde{M}_2^*)\widetilde{M}_2 + A_*^3\widetilde{M}_3 + \mathcal{O}(g^4) \\ &= A_* - iA_*^2(\widetilde{M}_2 - \widetilde{M}_2^*) + A_*^3(\widetilde{M}_3 - \widetilde{M}_3^* + 2\widetilde{M}_2^*(\widetilde{M}_2 - \widetilde{M}_2^*)) + \mathcal{O}(g^4). \end{aligned}$$

Therefore, for renormalizability, we need that as the cutoff is removed, both $\widetilde{M}_2 - \widetilde{M}_2^*$ and $\widetilde{M}_3 - \widetilde{M}_3^* + 2\widetilde{M}_2^*(\widetilde{M}_2 - \widetilde{M}_2^*)$ converge to finite limits. The first one we already know, and this is true for the second one also, but much more complicated. We will not attempt to prove it. We end with the remark that the convergence of a quantity as complicated as $\widetilde{M}_3 - \widetilde{M}_3^* + 2\widetilde{M}_2^*(\widetilde{M}_2 - \widetilde{M}_2^*)$ to a finite limit as the cutoff is removed indicates that something deeper is going on. Indeed, more is true: φ^4 theory is renormalizable in every order of perturbation theory. However, we do not have the time to prove that in this course.

22.5. Counterterms for two-loop renormalization

Instead of the relatively complicated derivation above, we could have also used the counterterm method for two-loop renormalization in φ^4 theory. The proper way to use the counterterm method for this purpose is to reparametrize g as

$$g = g_r + ig_r^2\widetilde{M}_2^* - g_r^3(\widetilde{M}_3^* + 2(\widetilde{M}_2^*)^2).$$

Indeed, an easy verification shows that plugging the above into (22.1) gives

$$A = g_r - ig_r^2(\widetilde{M}_2 - \widetilde{M}_2^*) + g_r^3(\widetilde{M}_3 - \widetilde{M}_3^* + 2\widetilde{M}_2^*(\widetilde{M}_2 - \widetilde{M}_2^*)) + \mathcal{O}(g_r^4).$$

Thus, if we accept out previous claim that $\widetilde{M}_3 - \widetilde{M}_3^* + 2\widetilde{M}_2^*(\widetilde{M}_2 - \widetilde{M}_2^*)$ converges to a finite limit as the cutoff is removed, then there are no infinities in the third order of perturbation theory for φ^4 theory after adding the counterterms $\frac{1}{4!}ig_r^2\widetilde{M}_2^*{:}\varphi(x)^4{:}$ and $-\frac{1}{4!}g_r^3(\widetilde{M}_3 - \widetilde{M}_3^* + 2\widetilde{M}_2^*(\widetilde{M}_2 - \widetilde{M}_2^*)){:}\varphi(x)^4{:}$ to the the interaction Hamiltonian density $\frac{1}{4!}g_r{:}\varphi(x)^4{:}$.

LECTURE 23

The model for free photons

Date: 11/14/2018
Scribe: Alex Dunlap

23.1. Photons

Photons are massless particles moving at the speed of light. In this lecture we will construct a model for free photons. This is a real model that is used in physics.

Recall that the four-momentum of a particle with mass m and velocity $\mathbf{v}$ (in some coordinate system) is the vector

$$p = \left(\frac{m}{\sqrt{1-\mathbf{v}^2}}, \frac{m\mathbf{v}}{\sqrt{1-\mathbf{v}^2}} \right). \tag{23.1}$$

For photons, $m = 0$ and $|\mathbf{v}| = 1$, so we get $0/0$ in (23.1). However, we also know that the four-momentum of a particle of mass m satisfies $p^2 = m^2$, so that $p \in X_m$. There is no problem in defining the manifold

$$X_0 = \{p : p^2 = 0, p^0 \geq 0\}.$$

(This is just the surface of a cone in $\mathbb{R}^{1,3}$.) So we can hypothesize that photons also have four-momenta, which belong to X_0. The manifold X_0 comes equipped with the Lorentz-invariant measure λ_0, defined as before. We will later use the fact that λ_0 has no point masses.

It is a physical fact the four-momentum is not the only quantity that characterizes a photon. A photon also has some internal degrees of freedom, called *spin*. Therefore it does not suffice to take $L^2(X_0, d\lambda_0)$ to be the Hilbert space for photons. Constructing the Hilbert space for photons is a slightly complicated affair, which we now carry out.

23.2. The Gupta–Bleuler construction

The procedure for constructing the Hilbert space for photons is called the *Gupta–Bleuler quantization*. It goes as follows. Start with the vector space of all $f : X_0 \to \mathbb{C}^4$ such that each coordinate of f is in $L^2(X_0, d\lambda_0)$, which is denoted by $L^2(X_0, d\lambda_0, \mathbb{C}^4)$. We will simply call this space $\mathcal{H}$. (This is *not* the final Hilbert space, but we'll use it.) For $w, z \in \mathbb{C}^4$, define the Minkowski inner product

$$(w, z) = (w^0)^* z^0 - (w^1)^* z^1 - (w^2)^* z^2 - (w^3)^* z^3.$$

If $\Lambda \in SO^{\uparrow}(1,3)$, define Λz as expected: if $z \in \mathbb{C}^4$ equals $x+iy$, with $x, y \in \mathbb{R}^4$, then let $\Lambda z = \Lambda x + i\Lambda y$. It is simple to check that $(\Lambda z, \Lambda w) = (z, w)$. For $\xi, \psi \in \mathcal{H}$, define

$$(\xi, \psi) = -\int d\lambda_0(p)\,(\xi(p), \psi(p)). \tag{23.2}$$

This defines a sesquilinear form on $\mathcal{H}$, but it is not positive-definite. To extract a subspace on which this form is positive semidefinite, define

$$\mathcal{H}' = \{\xi \in \mathcal{H} : (p, \xi(p)) = 0 \text{ for } \lambda_0\text{-a.e. } p\}.$$

LEMMA 23.1. *The inner product defined in* (23.2) *is positive semidefinite on* $\mathcal{H}'$.

PROOF. Take any $\xi \in \mathcal{H}'$. Suppose that p is a point so that $(p, \xi(p)) = 0$, and $p \neq 0$. Write

$$p = (p^0, \mathbf{p}) = (p^0, p^1, p^2, p^3).$$

Since $p \in X_0$, we have

$$(p_0)^2 = (p^1)^2 + (p^2)^2 + (p^3)^2.$$

Also, $p \neq 0$. So $p^0 > 0$ and $\mathbf{p} \neq 0$. Let $A \in SO(3)$ be such that $A\mathbf{p} = (0, 0, a)$ for some a. Since A is a rotation, $|a| = |\mathbf{p}|$. Moreover, we can arrange it so that $a = |\mathbf{p}| = p^0$. Let

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix} \in SO^{\uparrow}(1,3).$$

Let $x = \xi(p)$, and let $z = \Lambda x$. Note that $\Lambda p = (p^0, 0, 0, p^0)$. Thus,

$$(p, x) = (\Lambda p, \Lambda x) = (\Lambda p, z) = p^0(z^0 - z^3).$$

But we know that

$$(p, x) = (p, \xi(p)) = 0,$$

and $p^0 > 0$. Therefore, $z^0 = z^3$. Then

$$(\xi(p), \xi(p)) = (z, z) = (z^0)^2 - (z^1)^2 - (z^2)^2 - (z^3)^2 = -(z^1)^2 - (z^2)^2 \leq 0.$$

Therefore, $(\xi, \xi) \geq 0$, using that λ_0 has no point mass at 0. This completes the proof of the lemma. □

Having defined $\mathcal{H}'$, define $\mathcal{H}_{null} = \{\xi \in \mathcal{H}' : (\xi, \xi) = 0\}$. Define $\mathcal{H}_{phys} = \mathcal{H}'/\mathcal{H}_{null}$. Then $\mathcal{H}_{phys}$ is a Hilbert space under the Minkowski inner product. One can also check that it is complete. This is the Hilbert space for a single photon. For arbitrary numbers of photons, we take the bosonic Fock space of $\mathcal{H}_{phys}$.

23.3. The representation for a free photon

Having defined the Hilbert space, it remains to define the representation of the Poincaré group that describes the spacetime trajectory of a photon. Recall that in quantum field theory, the full spacetime description of a system is given by a state in a Hilbert space. Recall also the definition of the Poincaré group $\mathcal{P} = \mathbb{R}^{1,3} \rtimes SO^\uparrow(1,3)$, which is the group of symmetries of $\mathbb{R}^{1,3}$. The action is of $(a, A) \in \mathcal{P}$ on a point $x \in \mathbb{R}^{1,3}$ is $(a, A)x = a + Ax$. Finally, recall the postulate that for any quantum system, there is a strongly continuous unitary representation U of $\mathcal{P}$ in the Hilbert space for the system, such that if an observer sees the system in state ψ, then in the different coordinate system obtained by the action of (a, A), the system appears to be in state $U(a, A)\psi$.

The representation U of $\mathcal{P}$ in $\mathcal{H}_{phys}$ that describes the evolution of a free photon is defined as follows. For $\psi \in \mathcal{H}$, define

$$(U(a, A)\psi)(p) = e^{i(a,p)} A\psi(A^{-1}p).$$

It is easy to check that this is a representation of $\mathcal{P}$ in $\mathcal{H}$. Moreover, it is easy to check that $U(a, A)$ maps $\mathcal{H}'$ into $\mathcal{H}'$, and moreover it descends to a representation in $\mathcal{H}_{phys}$. To prove the first claim, note that since λ_0 is Lorentz-invariant and $(p, \psi(p)) = 0$ for λ_0-a.e. p for any $\psi \in \mathcal{H}'$,

$$\begin{aligned}
\int d\lambda_0(p)\, (p, U(a, A)\psi(p)) &= \int d\lambda_0(p)\, (p, e^{i(a,p)} A\psi(A^{-1}p)) \\
&= \int d\lambda_0(p)\, e^{i(a,p)} \underbrace{(A^{-1}p, \psi(A^{-1}p))}_{= 0 \text{ for } \lambda_0\text{-a.e. } p} \\
&= 0.
\end{aligned}$$

The proof of the second claim follows by showing that $U(a, A)$ maps $\mathcal{H}_{null}$ into $\mathcal{H}_{null}$ in a similar manner. Thus, U is a representation of $\mathcal{P}$ in $\mathcal{H}_{phys}$. Moreover, U is a unitary representation:

$$\begin{aligned}
(U(a, A)\psi, U(a, A)\xi) &= -\int d\lambda_0(p)\, (e^{i(a,p)} A\psi(A^{-1}p), e^{i(a,p)} A\xi(A^{-1}p)) \\
&= -\int d\lambda_0(p)\, (A\psi(A^{-1}p), A\xi(A^{-1}p)) \\
&= -\int d\lambda_0(p)\, (\psi(A^{-1}p), \xi(A^{-1}p)) \\
&= -\int d\lambda_0(p)\, (\psi(p), \xi(p)) \\
&= (\psi, \xi),
\end{aligned}$$

where in the third equality, we used the fact that the inner product is Lorentz-invariant, and in the fourth inequality, we used the fact that λ_0 is Lorentz-invariant.

23.4. The quantized electromagnetic four-potential

Just as we had the massive free field associated with our massive scalar boson, there is a field associated with photons. It's called the electromagnetic four-potential. Consider our original $\mathcal{H} = L^2(X_0, d\lambda_0, \mathbb{C}^4)$ with the Euclidean inner product, defined as

$$(\xi, \psi)_E := \int d\lambda_0(p) \sum_{\mu=0}^{3} \xi^\mu(p)^* \psi^\mu(p).$$

Then $\mathcal{H}$ is a Hilbert space under this inner product. Consider the Fock space $\mathcal{B}$ of this Hilbert space, and our familiar operator-valued distributions A and $A^\dagger$. Define two other operators B and $B^\dagger$ as follows. Recall the matrix η defined in (9.1). Then for any $\xi = (\xi^0, \xi^1, \xi^2, \xi^3) \in \mathcal{H}$, we can naturally define $\eta\xi = (\xi^0, -\xi^1, -\xi^2, -\xi^3)$. Let

$$B(\xi) = A(-\eta\xi),$$
$$B^\dagger(\xi) = A^\dagger(\xi).$$

Then B, $B^\dagger$ are also operator-valued distributions on $\mathcal{H}$, and they have the properties that

$$[B(\xi), B(\psi)] = [B^\dagger(\xi), B^\dagger(\psi)] = 0,$$
$$[B(\xi), B^\dagger(\psi)] = \underbrace{(\xi, \psi)}_{\text{our inner product}} 1.$$

This is because

$$[B(\xi), B^\dagger(\psi)] = [A(-\eta\xi), A^\dagger(\psi)] = (-\eta\xi, \psi)_E 1 = (\xi, \psi)1.$$

For $f \in \mathscr{S}(\mathbb{R}^4)$, $\mu = 0, 1, 2, 3$, define

$$A_\mu(f) = B(\widehat{f} e_\mu) + B^\dagger(\widehat{f} e_\mu),$$

where e_μ is the μth standard basis vector and $\widehat{f}$ is the Fourier transform of f on $\mathbb{R}^{1,3}$, restricted to X_0. The 4-tuple $A = (A_0, A_1, A_2, A_3)$ is called the quantized electromagnetic four-potential. Each A_μ is an operator-valued distribution that takes $f \in \mathscr{S}(\mathbb{R}^4)$ to a linear operator $A_\mu(f)$ on $\mathcal{B}_0$, where $\mathcal{B}_0$ is our familiar dense subspace of the Fock space $\mathcal{B}$.

LECTURE 24

The electromagnetic field

Date: 11/26/2018
Scribe: Jae Hee Lee

24.1. Relationship between the two Hilbert spaces

Recall the spaces $\mathcal{H}$, $\mathcal{H}'$ and $\mathcal{H}_{phys}$ from the previous lecture. The space $\mathcal{H}_{phys}$ is the Hilbert space for photons, whereas the quantized electromagnetic four-potential A was defined on $\mathcal{H}$. Recall that $\mathcal{H}$ carries the Euclidean inner product, while the inner product on $\mathcal{H}_{phys}$ is derived from the Minkowski inner product on $\mathcal{H}'$. So how are the two Hilbert spaces related? More importantly, how does the electromagnetic four-potential act on the bosonic Fock space for photons? This happens as follows.

Let $\mathcal{B}$ be the bosonic Fock space of $\mathcal{H}$ and let $\mathcal{B}'$ be the bosonic Fock space of $\mathcal{H}'$ (all under the Euclidean inner product). Then $\mathcal{B}'$ is a closed subspace of $\mathcal{B}$. The Minkowski sesquilinear form and the Minkowski norm on $\mathcal{H}'$ extends to $\mathcal{B}'$ in the following way. We first define on the simple tensors

$$(\xi_1 \otimes \cdots \otimes \xi_n, \psi_1 \otimes \cdots \otimes \xi_n) = \prod_{i=1}^{n} (\xi_i, \psi_i),$$

and use linearity and the inequality

$$\|\xi\| = \sqrt{(\xi, \xi)} \leq \|\xi\|_E \quad \text{for } \xi \in \mathcal{H}'$$

to extend this to $\mathcal{B}'$. Note that the expression $\sqrt{(\xi, \xi)}$ makes sense because the Minkowski form is nonnegative on $\mathcal{H}'$. Next, we define

$$\mathcal{B}_{null} := \{\psi \in \mathcal{B}' : \|\psi\| = 0\},$$

which is a closed subspace of $\mathcal{B}'$, and correspondingly define

$$\mathcal{B}_{phys} = \mathcal{B}'/\mathcal{B}_{null}.$$

Note that $\mathcal{B}_{phys}$ is equipped with the Euclidean inner product.

One can show that there is a canonical Hilbert space isomorphism between $\mathcal{B}_{phys}$ and the bosonic Fock space of $\mathcal{H}_{phys}$ (which was constructed using the Minkowski inner product). The construction of this isomorphism, however, is slightly involved.

The upshot is that we can now completely forget about the Minkowski norm and always work with $\mathcal{B}_{phys}$ with the Euclidean inner product as the

bosonic Fock space for photons. In future calculations for quantum electrodynamics, we will simply work with the bosonic Fock space $\mathcal{B}$ of the Hilbert space $\mathcal{H}$, and the operator-valued distributions $(A_\mu)_{0\leq\mu\leq 3}$ acting on this space.

24.2. Action of the electromagnetic four-potential

Recall that for $f \in \mathscr{S}(\mathbb{R}^4)$ and $0 \leq \mu \leq 3$, the operator $A_\mu(f)$ acts on $\mathcal{B}$. However, it is not clear if it defines an operator on the quotient space $\mathcal{B}_{phys}$. In particular this is because we don't know if $\mathcal{B}_{null}$ gets sent into $\mathcal{B}_{null}$ by these operators.

In fact, this is *not true* for the A_μ's. What is true, yet, is that certain linear combinations of A_μ's are well-defined on $\mathcal{B}_{phys}$. For example, define

$$F_{\mu\nu} := \partial_\mu A_\nu - \partial_\nu A_\mu,$$

where ∂_μ denotes the partial derivative $\partial/\partial x^\mu$. Then it turns out that for any $f \in \mathscr{S}(\mathbb{R}^4)$, $F_{\mu\nu}(f)$ is a well-defined operator on $\mathcal{B}_{\text{phys}}$. The array $F = (F_{\mu\nu})_{0\leq\mu,\nu\leq 3}$ of operator-valued distributions is known as the **electromagnetic field**, for reasons that will be clear in the next section.

Another example of a class of operators that act on $\mathcal{B}_{phys}$ is the following. Suppose that $f_0, f_1, f_2, f_3 \in \mathscr{S}(\mathbb{R}^4)$ satisfy

$$\partial_0 f_0 - \partial_1 f_1 - \partial_2 f_2 - \partial_3 f_3 \equiv 0.$$

Then it can be shown that

$$A_0(f_0) - A_1(f_1) - A_2(f_2) - A_3(f_3) \tag{24.1}$$

acts on $\mathcal{B}_{phys}$.

The physical meaning of the above discussion is the following. The electromagnetic four-potential A is not directly observable; it is only an abstract object. The things that we can observe are exactly those operators that are defined on $\mathcal{B}_{phys}$.

24.3. Classical (Maxwell) theory of electromagnetism

Let us recall the classical theory of electromagnetism. In Maxwell's formulation, we are given two time-dependent vector fields on our physical space, the *electric field*

$$\mathbf{E} : \mathbb{R}^{1,3} \to \mathbb{R}^3$$

and the *magnetic field*

$$\mathbf{B} : \mathbb{R}^{1,3} \to \mathbb{R}^3.$$

The evolution of these fields is governed by the Maxwell equations, which we write in the units where $c = 1$:

$$\nabla \cdot \mathbf{E} = 4\pi\rho \tag{24.2}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{24.3}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{24.4}$$

$$\nabla \times \mathbf{B} = 4\pi\mathbf{J} + \frac{\partial \mathbf{E}}{\partial t} \tag{24.5}$$

where ρ is the charge density (charge per unit volume) and $\mathbf{J}$ is the current density (current per unit area). In particular by knowing both ρ and $\mathbf{J}$ we may determine the evolution of the electric and magnetic fields.

In the absence of electrons, we have $\rho \equiv 0$ and $\mathbf{J} \equiv 0$. The equations that we get in this situation are known as 'Maxwell's equations in the vacuum'. A class of solutions of Maxwell's equations in the vacuum can be constructed as follows. Take any smooth vector field

$$A : \mathbb{R}^{1,3} \to \mathbb{R}^4, \quad A = (A_0, A_1, A_2, A_3)$$

and define the *classical electromagnetic field* $F = (F_{\mu\nu})_{0\leq\mu,\nu\leq 3}$ as

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu.$$

Given F, define two vector fields $\mathbf{E}$ and $\mathbf{B}$ as

$$F = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}.$$

Then from the definition of F, one gets the "Bianchi identity"

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0 \quad \forall \lambda, \mu, \nu$$

which implies equation (24.3) and (24.4) of the Maxwell equations in the vacuum. If we further assume that A satisfies the "equation of evolution"

$$\partial_0 F_{0\nu} - \partial_1 F_{1\nu} - \partial_2 F_{2\nu} - \partial_3 F_{3\nu} = 0 \quad \forall \nu. \tag{24.6}$$

we recover equations (24.2) and (24.5) of the Maxwell equations in the vacuum.

Now it turns out that *any solution of Maxwell's equations in the vacuum is representable in the above manner.* The vector field A is known as the classical electromagnetic four-potential. However, given $\mathbf{E}$ and $\mathbf{B}$, the field A is not unique; for any smooth $\lambda : \mathbb{R}^{1,3} \to \mathbb{R}$ (a gauge transformation),

$$A'_\mu(x) = A_\mu(x) + \partial_\mu \lambda(x)$$

also yields the same $\mathbf{E}$ and $\mathbf{B}$. In fact, this describes the set of all solutions for given $\mathbf{E}$ and $\mathbf{B}$. The procedure for choosing a unique representative is

called *gauge fixing*. There are many ways to do this, but one way is to impose the *Lorenz gauge condition*,

$$\partial_0 A_0 - \partial_1 A_1 - \partial_2 A_2 - \partial_3 A_3 \equiv 0.$$

(Note that this 'Lorenz' without a 't' — not the same as the Lorentz in Lorentz transforms.)

24.4. From classical to quantized

The quantized electromagnetic four-potential $A = (A_\mu)_{0 \le \mu \le 3}$ that we defined using the Gupta–Bleuler approach is a quantization of the classical electromagnetic four-potential, in the sense that it satisfies the vacuum Maxwell equations formally. Moreover, it also satisfies the Lorenz gauge condition, in the sense that for any $f \in \mathscr{S}(\mathbb{R}^4)$,

$$\partial_0 A_0(f) - \partial_1 A_1(f) - \partial_2 A_2(f) - \partial_3 A_3(f) = 0$$

as an operator on $\mathcal{B}_{phys}$.

24.5. A note on the Einstein convention

Many of the expressions in this lecture could have been written much more compactly using Einstein's conventions for special relativity. In particular, there are two important rules:

- If we have a 4-tuple of objects $(a_\mu)_{0 \le \mu \le 3}$, then raising μ changes the sign of the last three. That is, $a^0 = a_0$, but $a^1 = -a_1$, $a^2 = -a_2$, and $a^3 = -a_3$. If the indices were raised to begin with, the lowering them has the same effect. For example, we can raise the indices on ∂_μ to get ∂^μ.
- If an index is repeated within the same term, that indicates that we have to sum over that index. For example,

$$\partial^\mu A_\mu = \sum_{\mu=0}^{3} \partial^\mu A_\mu = \partial_0 A_0 - \partial_1 A_1 - \partial_2 A_2 - \partial_3 A_3.$$

Thus, for example, the Lorenz gauge condition can be compactly written as $\partial^\mu A_\mu = 0$. Similarly, the equation of motion (24.6) can be written as $\partial^\mu F_{\mu\nu} = 0$ for each ν. The expression (24.1) becomes $A_\mu(f^\mu)$. The Minkowski inner product between $x, y \in \mathbb{R}^{1,3}$ is $x^\mu y_\mu$.

LECTURE 25

The model for free electrons

Date: 11/28/2018
Scribe: Kevin Yang

25.1. Projective representations

Today, we will construct a model for electrons that also applies for positrons and several other types of particles. Let m equal the rest mass of the electron. Let $\mathcal{H} = L^2(X_m, d\lambda_m, \mathbb{C}^2)$ be the space of $\mathbb{C}^2$-valued L^2-functions on the manifold X_m under the measure $d\lambda_m$. The spacetime state of an electron is described by an element of this space. There are two basic questions to answer before we proceed any further:

(1) What is the inner product on this space?
(2) What is the representation of the Poincaré group, $\mathcal{P}$?

To answer the second question, we need to modify the fifth postulate of quantum field theory stated in Lecture 11. Recall that this postulate states that if a system looks like it is in state $\psi \in \mathcal{H}$ in some choice of coordinate system, then upon changing the coordinate system by some $(a, A) \in \mathcal{P}$, the system looks like it is in the state $U(a, A)\psi$, where U is some unitary representation of $\mathcal{P}$ in $\mathcal{H}$. The representation property ensures that $U((a, A)(b, B)) = U(a, A)U(b, B)$, meaning that two successive changes of coordinate system amount to one composite change, and unitarity ensures that normalized states are sent to normalized states (conservation of probability). Now note that expectations of observables under a state ψ are invariant under multiplication of the state by a phase $\alpha \in \mathbb{C}$, $|\alpha| = 1$. Thus, we do not really need U two a unitary representation; it suffices that U is unitary and is a *projective representation*, meaning that

$$U((a, A), (b, B)) = r(a, A, b, B)U(a, A)U(b, B),$$

where $r(a, A, b, B) \in \mathbb{C}$ and $|r(a, A, b, B)| = 1$.

25.2. The representation for electrons

Let $SL(2, \mathbb{C})$ be the group of 2×2 complex matrices of determinant 1. For any point $x \in \mathbb{R}^{1,3}$, define the matrix

$$M(x) = \begin{pmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{pmatrix}.$$

One can check the above definition provides a bijection between $\mathbb{R}^{1,3}$ and the space of 2×2-complex Hermitian matrices. Moreover,

$$\det M(x) = (x, x),$$

where the bracket on the right is the Minkowski inner product.

Now take any $A \in SL(2, \mathbb{C})$. Then for any $x \in \mathbb{R}^{1,3}$, the matrix $AM(x)A^\dagger$ is again a 2×2 complex Hermitian matrix. Define $\kappa(A)(x)$ to be the unique $y \in \mathbb{R}^{1,3}$ such that

$$M(y) = AM(x)A^\dagger.$$

The following facts about κ are not hard to prove:

- For any $A \in SL(2, \mathbb{C})$, $\kappa(A)$ is a linear map on $\mathbb{R}^{1,3}$.
- κ is a 2-to-1 surjective homomorphism from $SL(2, \mathbb{C})$ onto $SO^\uparrow(1, 3)$.
- $\kappa(A) = \kappa(B)$ if and only if either $A = B$ or $A = -B$. (We will henceforth write this as $A = \pm B$.)
- $\kappa(A^\dagger) = \kappa(A)^\dagger$, where $M^\dagger$ denotes the conjugate transpose of a matrix M.

By the second property listed above, there are many maps $\rho : SO^\uparrow(1, 3) \to SL(2, \mathbb{C})$ such that $\kappa \circ \varrho$ is the identity on $SO^\uparrow(1, 3)$. Moreover, any such ϱ satisfies $\varrho(AB) = \pm\varrho(A)\varrho(B)$. It turns out that due to algebraic constraints, we cannot choose ϱ to be a homomorphism. The fourth property implies that $\varrho(A^\dagger) = \pm\varrho(A)^\dagger$.

25.3. Constructing the inner product

We assume that nature provides a ϱ as above so that the projective representation of $SO^\uparrow(1, 3)$ in the space $\mathcal{H} = L^2(X_m, d\lambda_m, \mathbb{C}^2)$ is given by

$$(U(a, A)\psi)(p) \;=\; \mathrm{e}^{i(a,p)}\varrho(A)\psi(A^{-1}p).$$

One can check that this is inded a projective representation. To prove that the above representation is unitary, we first have to define the inner product. We now build the inner product on $\mathcal{H}$. Let $p^* = (m, 0, 0, 0) \in X_m$ denote a distinguished element of the manifold X_m. The action of $SO^\uparrow(1, 3)$ on X_m is transitive, so for any $p \in X_m$ there are maps taking p^* to p. If A, B are any such maps, then they are related by a spatial rotation:

$$A = BR, \quad R = \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix}, \quad Q \in SO(3).$$

In words, R fixes time and acts by rotation on the spatial variables. It turns out that there exists a standard way to choose a unique element from the class of all such maps taking p^* to p. Precisely, for any $p \in X_m$ there exists a unique positive-definite $V_p \in \mathrm{SL}(2, \mathbb{C})$ such that $\kappa(V_p)(p^*) = p$. The uniqueness is provided by the positive-definite constraint. In the physics literature, the map $\kappa(V_p)$ is known as the "pure boost" that takes p^* to p.

With this machinery in place, we may now define the following inner product on $\mathcal{H}$:

$$(\psi, \varphi) := \int_{X_m} d\lambda_m(p)\psi(p)^\dagger V_p^{-2}\varphi(p), \quad \psi, \varphi \in \mathrm{L}^2(X_m, d\lambda_m, \mathbb{C}^2).$$

It is not difficult to verify that $\mathcal{H}$ is indeed a Hilbert space under this inner product. We now need to check that, under this inner product, the projective representation U is unitary. To this end, we first note

$$(U(a,A)\psi)(p)^\dagger V_p^{-2}(U(a,A)\varphi)(p) \;=\; \psi(A^{-1}p)^\dagger \varrho(A)^\dagger V_p^{-2}\varrho(A)\varphi(A^{-1}p).$$

Integrating the right side with respect to $d\lambda_m$ on X_m, we get

$$(U(a,A)\psi, U(a,A)\varphi) = \int_{X_m} d\lambda_m(p)\psi(p)^\dagger \varrho(A)^\dagger V_{Ap}^{-2}\varrho(A)\varphi(p)$$

by the Lorentz invariance of $d\lambda_m$. It now remains to prove the following identity.

LEMMA 25.1. *For any A and p,*

$$\varrho(A)^\dagger V_{Ap}^{-2}\varrho(A) \;=\; V_p^{-2}.$$

PROOF. Since κ is a homomorphism, we know

$$\kappa(V_{Ap}^{-1}\varrho(A)) = \kappa(V_{Ap}^{-1})\kappa(\varrho(A)) = \kappa(V_{Ap})^{-1}A.$$

By definition of V, we know

$$A^{-1}\kappa(V_{Ap})p^* = A^{-1}Ap = p.$$

Thus, $A^{-1}\kappa(V_{Ap}) = \kappa(V_p)R$ for some spatial rotation R. Applying ϱ to both sides, we get

$$\varrho(A)^{-1}V_{Ap} = \pm V_p\varrho(R).$$

This, in turn, implies

$$\varrho(A)^\dagger V_{Ap}^{-2}\varrho(A) = V_p^{-1}(\varrho(R)^{-1})^\dagger\varrho(R)^{-1}V_p^{-1} = \pm V_p^{-2},$$

since $\varrho(R)\varrho(R)^\dagger = \pm\varrho(RR^\dagger) = \pm I$. But $\varrho(A)^\dagger V_{Ap}^{-2}\varrho(A)$ and V_p^{-2} are both positive definite matrices, and therefore they must be equal. This completes the proof. □

25.4. Fermionic Fock Space

We have now built the state space for a single electron. Because electrons are fermionic in nature, to build the state space for n electrons we need to build a fermonic Fock space instead of the bosonic Fock space. This is defined as follows.

Let $\mathcal{H}$ denote the Hilbert space for a single fermionic particle. Let $\{e_j\}_{j=1}^{\infty}$ be an orthonormal basis of $\mathcal{H}$. The appropriate state space for n such particles is

$$\mathcal{H}_{anti}^{\otimes n} = \Bigg\{ \sum_{i_1,\ldots,i_n \geq 1} \alpha_{i_1,\ldots,i_n} e_{i_1} \otimes \ldots \otimes e_{i_n} : \ \forall \sigma \in S_n, \\ \alpha_{i_{\sigma(1)},\ldots,i_{\sigma(n)}} = \operatorname{sign}(\sigma) \alpha_{i_1,\ldots,i_n} \Bigg\}.$$

With $\alpha_{i_1,\ldots,i_n}$ as above, note that $\alpha_{i_1,\ldots,i_n} = 0$ if $i_1, \ldots, i_n$ are not distinct indices. Indeed, if we had a pair of repeated indices, the transposition swapping their positions would preserve the index-tuple. Since transpositions are odd, this would imply $\alpha_{i_1,\ldots,i_n} = -\alpha_{i_1,\ldots,i_n}$.

Like the bosonic versions, these spaces are actually basis-independent. We now define

$$\mathcal{F}_0 = \bigoplus_{n=0}^{\infty} \mathcal{H}_{anti}^{\otimes n},$$

and let $\mathcal{F}$ denote the closure of $\mathcal{F}_0$ under the natural inner product. In the context of electrons, we will take $\mathcal{H} = \mathrm{L}^2(X_m, d\lambda_m, \mathbb{C}^2)$, and the state space for arbitrary numbers of electrons to be the fermionic Fock space of $\mathcal{H}$.

We conclude with a simple example. If $\mathcal{H} = L^2(\mathbb{R})$, then the fermionic Fock space of n particles is

$$\mathcal{H}_{anti}^{\otimes n} = \Big\{ f \in L^2(\mathbb{R}^n) : \ f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = \operatorname{sign}(\sigma) f(x_1, \ldots, x_n), \\ \forall x \in \mathbb{R}^n, \forall \sigma \in S_n \Big\}.$$

LECTURE 26

The Dirac field

Date: 11/30/2018
Scribe: Cole Graham

26.1. A basis for fermionic Fock space

Given an orthonormal basis $e_1, e_2, \ldots$ of $\mathcal{H}$, we can form an orthonormal basis for $\mathcal{H}_{anti}^{\otimes n}$ from elements of the form

$$\frac{1}{\sqrt{n!}} \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) e_{i_{\sigma(1)}} \otimes \cdots \otimes e_{i_{\sigma(n)}},$$

where $i_1 < i_2 < \cdots < i_n$. We denote this element by $|n_1, n_2, \ldots\rangle$, where

$$n_j = \begin{cases} 1 & \text{if } j \in \{i_1, \ldots, i_n\}, \\ 0 & \text{otherwise.} \end{cases}$$

We may interpret the multi-particle state $|n_1, \ldots\rangle$ as n particles in states $e_{i_1}, \ldots, e_{i_n}$. As remarked in the previous lecture, the antisymmetry forbids occupation numbers above 1. This is the *Pauli exclusion principle*: two identical fermions cannot simultaneously occupy the same state.

From the above discussion, we have that

$$\left\{ |n_1, n_2, \ldots\rangle \,:\, n_i \in \{0, 1\} \text{ and } \sum_i n_i = n \right\}$$

is an orthonormal basis for $\mathcal{H}_{anti}^{\otimes n}$. Allowing n to vary in $\mathbb{Z}_{\geq 0}$, we obtain an orthonormal basis for the fermionic Fock space $\mathcal{F}$:

$$\left\{ |n_1, n_2, \ldots\rangle \,:\, n_i \in \{0, 1\} \text{ and } \sum_i n_i < \infty \right\}.$$

26.2. Creation and annihilation operators on $\mathcal{F}$

We begin by defining creation and annihilation operators relative to a fixed orthonormal basis $\{e_i\}$ of $\mathcal{H}$. For each $k \geq 1$, define the creation operator $a_k^\dagger$ as follows. Let $\psi = |n_1, n_2, \ldots\rangle$. If $n_k = 1$, let $a_k^\dagger \psi = 0$, where 0 denotes the zero of the vector space $\mathcal{F}$, not the vacuum state. If instead $n_k = 0$, let

$$a_k^\dagger \psi = (-1)^m \, |n_1, \ldots, n_{k-1}, 1, n_{k+1}, \ldots\rangle \,,$$

where

$$m \coloneqq \#\{1 \le i < k \,:\, n_i = 1\} = \sum_{i=1}^{k-1} n_i.$$

REMARK 26.1. If $\psi = |0, 0, \ldots\rangle$ is the vacuum state, $a_k^\dagger \psi$ is the state with one particle in state e_k, as expected.

We also have an annihilation operator a_k. If $n_k = 0$, let $a_k\psi = 0$. If instead $n_k = 1$, let

$$a_k\psi = (-1)^m \left|n_1, \ldots, n_{k-1}, 0, n_{k+1}, \ldots\right\rangle,$$

with m as above. The sign correction $(-1)^m$ will ensure simpler formulæ in the sequel.

We now define continuous versions of $a_k^\dagger$ and a_k. We want maps $A^\dagger$ and A from $\mathcal{H}$ into $\mathscr{L}(\mathcal{F}_0, \mathcal{F})$, where $\mathscr{L}(\mathcal{F}_0, \mathcal{F})$ denotes the space of linear maps from $\mathcal{F}_0$ to $\mathcal{F}$. Given $f \in \mathcal{H}$, write

$$f = \sum_{k=1}^{\infty} \alpha_k e_k$$

and define

$$A^\dagger(f)\psi \coloneqq \sum_k \alpha_k a_k^\dagger \psi, \quad A(f)\psi \coloneqq \sum_k \alpha_k^* a_k \psi.$$

REMARK 26.2. Although we do not show it here, $A^\dagger, A$ are basis independent. That is, they do not depend on the choice of basis $\{e_i\}$ for $\mathcal{H}$.

We can easily check the following *anticommutation* relations:

$$\{a_k, a_\ell^\dagger\} = \delta_{k,\ell} 1, \quad \{a_k^\dagger, a_\ell^\dagger\} = 0, \quad \{a_k, a_\ell\} = 0 \quad \text{for all } k, \ell \in \mathbb{N}.$$

Here $\{A, B\} \coloneqq AB + BA$ is the anticommutator of A and B. These relations imply

$$\{A(\xi), A^\dagger(\eta)\} = (\xi, \eta) 1, \quad \{A^\dagger(\xi), A^\dagger(\eta)\} = 0, \quad \{A(\xi), A(\eta)\} = 0$$

for all $\xi, \eta \in \mathcal{H}$.

Now if U is a unitary operator on $\mathcal{H}$, it has a natural extension to $\mathcal{F}$ (by acting on each factor in an elementary tensor). As in the bosonic case, we have

$$UA(\xi)U^{-1} = A(U\xi) \quad \text{and} \quad UA^\dagger(\xi)U^{-1} = A^\dagger(U\xi) \quad \text{for all } \xi \in \mathcal{H}.$$

An analog of Wick's theorem holds in the fermionic case, with one wrinkle: the resulting sum is signed.

26.3. The Dirac field

Recall that massive scalar bosons have a corresponding free field, and photons the electromagnetic field. We now define the *Dirac field* associated to electrons (and all spin-$\frac{1}{2}$ particles).

Let m denote the mass of the electron. In the last lecture, we defined the single-electron state-space $\mathcal{H} = L^2(X_m, \mathrm{d}\lambda_m, \mathbb{C}^2)$, and endowed it with an inner product based on our projective representation of the Poincaré group. Let $\mathcal{F}$ denote the fermionic Fock space corresponding to $\mathcal{H}$, and let e_1, e_2 denote the standard basis vectors of $\mathbb{C}^2$.

We will formally define operator-valued distributions $a^\dagger(p,s)$ and $a(p,s)$ for $p \in X_m$ and $s \in \{1,2\}$. For $\xi \in L^2(X_m, \mathrm{d}\lambda_m)$, we say

$$A^\dagger(\xi e_s) = \int_{X_m} d\lambda_m(p)\ \xi(p) a^\dagger(p,s)$$
$$A(\xi e_s) = \int_{X_m} d\lambda_m(p)\ \xi(p)^* a(p,s).$$

Thus formally

$$a^\dagger(p,s) = A^\dagger(\delta_p e_s) \quad \text{and} \quad a(p,s) = A(\delta_p e_s).$$

This is wholly analogous to the case of scalar bosons.

Recall that X_m is parameterized by $\mathbf{p} \in \mathbb{R}^3$ with $p = (p^0, \mathbf{p})$ and $p^0 = \omega_{\mathbf{p}} \coloneqq \sqrt{m^2 + \mathbf{p}^2}$. We define

$$a^\dagger(\mathbf{p}, s) \coloneqq \frac{a^\dagger(p,s)}{\sqrt{2\omega_{\mathbf{p}}}} \quad \text{and} \quad a(\mathbf{p}, s) \coloneqq \frac{a(p,s)}{\sqrt{2\omega_{\mathbf{p}}}}.$$

Then we can check that

$$\{a(\mathbf{p}, s), a^\dagger(\mathbf{p}', s')\} = (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{p}') \delta_{s,s'} 1,$$

and all other anticommutators vanish. Again, these are much like the commutator relations in the scalar boson case.

Now recall our distinguished element $p^* = (m, 0, 0, 0) \in X_m$. For each $p \in X_m$, recall that there is a unique positive-definite $V_p \in SL(2, \mathbb{C})$ such that

$$\kappa(V_p) p^* = p.$$

Here $\kappa \colon SL(2,\mathbb{C}) \to SO^\uparrow(1,3)$ denotes the double-cover defined in the previous lecture. The transformation $\kappa(V_p)$ is the "pure boost" that takes p^* to p. Define $S_p \in GL(4, \mathbb{C})$ by

$$S_p \coloneqq \begin{pmatrix} (V_p^\dagger)^{-1} & 0 \\ 0 & V_p \end{pmatrix}.$$

Also, define a basis $\{f_1, f_2, f_1', f_2'\}$ of $\mathbb{C}^4$:

$$f_1 \coloneqq \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad f_2 \coloneqq \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad f_1' \coloneqq \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \quad f_2' \coloneqq \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}.$$

Finally, define

$$u(\mathbf{p}, s) \coloneqq S_p f_s \quad \text{and} \quad v(\mathbf{p}, s) \coloneqq S_p f_s'. \tag{26.1}$$

Let u_k, v_k denote the components of u, v for $k \in \{1, 2, 3, 4\}$.

We have defined $\mathcal{F}$ to be the fermionic Fock space for electrons. Construct an identical copy $\mathcal{F}'$ for positrons. Then let $\mathcal{G} \coloneqq \mathcal{F} \otimes \mathcal{F}'$ denote the joint electron-positron space. Let $b^\dagger$ and b denote the analogs of $a^\dagger$ and a on $\mathscr{F}'$. Note that any operator C on $\mathcal{F}$ extends uniquely to $\mathcal{G}$ via

$$C(f \otimes f') \coloneqq Cf \otimes f'.$$

Operators likewise extend from $\mathcal{F}'$ to $\mathcal{G}$.

For each $k \in \{1, 2, 3, 4\}$, we formally define an operator-valued distribution ψ_k on $\mathbb{R}^{1,3}$ as

$$\psi_k(x) \coloneqq \sum_{s=1}^{2} \int_{\mathbb{R}^3} \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{\sqrt{m}}{\sqrt{2\omega_{\mathbf{p}}}} \Big[e^{-i(x,p)} u_k(\mathbf{p}, s) a(\mathbf{p}, s) + e^{i(x,p)} v_k(\mathbf{p}, s) b^\dagger(\mathbf{p}, s) \Big]. \tag{26.2}$$

We then form the 4-tuple of operator valued distributions

$$\psi(x) \coloneqq \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \\ \psi_3(x) \\ \psi_4(x) \end{pmatrix}.$$

We should consider how ψ differs from the free field φ. Of course ψ has four components, sums over the two spins $s = 1, 2$, and involves factors u and v from the projective representation. Most importantly, however, ψ involves a and $b^\dagger$ rather than simply a and $a^\dagger$. Early attempts to treat electrons ran into inconsistencies when using just $\mathcal{F}$. Dirac had the insight to add an extra Fock space $\mathcal{F}'$, which resolved all difficulties in the theory. Indeed, the strength of the theory led Dirac to *predict* the existence of the positron before its experimental discovery.

LECTURE 27

Introduction to quantum electrodynamics

Date: 12/3/2018
Scribe: Yang Liu

27.1. Pauli and Dirac matrices

The Pauli matrices are the following four 2×2 matrices:

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\ \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},\ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The Dirac matrices, defined using Pauli matrices, are the following four 4×4 matrices:

$$\gamma^0 = \begin{pmatrix} \sigma_0 & 0 \\ 0 & -\sigma_0 \end{pmatrix},\ \gamma^1 = \begin{pmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{pmatrix},\ \gamma^2 = \begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{pmatrix},\ \gamma^3 = \begin{pmatrix} 0 & \sigma_3 \\ -\sigma_3 & 0 \end{pmatrix}.$$

The *Dirac adjoint* of any column vector $x \in \mathbb{R}^4$ is the row vector

$$\overline{x} := x^\dagger \gamma^0,$$

where $x^\dagger$ is the conjugate transpose of x. The definition extends to the Dirac field ψ, whose Dirac adjoint $\overline{\psi}$ is defined as

$$\overline{\psi} := \psi^\dagger \gamma^0.$$

Note that $\overline{\psi}$ is a 1×4 row vector of operator valued distributions.

27.2. The interaction Hamiltonian for QED

Recall the fermionic Fock spaces $\mathcal{F}$ and $\mathcal{F}'$ for electrons and positrons, and the projective unitary representation of the Poincaré group associated with free evolution on these Fock spaces. Recall also the Hilbert space $\mathcal{H} = L^2(X_0, d\lambda_0, \mathbb{C}^4)$ with Euclidean inner product, and its bosonic Fock space $\mathcal{B}$, on which we defined the electromagnetic four-potential $A = (A_\mu)_{0 \le \mu \le 3}$. Finally, recall the unitary representation of the Poincaré group on this space that was used to describe free evolution of photons.

Let $\mathcal{R} = \mathcal{B} \otimes \mathcal{F} \otimes \mathcal{F}'$. The above discussion implies that there is a natural free evolution Hamiltonian H_0 on this tensor product space, which represents free evolution in each of the component spaces. Quantum electrodynamics, which is the model of photons, electrons and positrons evolving in time by

interacting with each other, is described by a Hamiltonian $H = H_0 + H_I$, where H_I is the following interaction Hamiltonian:

$$H_I = e \sum_{\mu=0}^{3} \int d^3\mathbf{x} :\overline{\psi}(0,\mathbf{x})\gamma^\mu \psi(0,\mathbf{x}) A_\mu(0,\mathbf{x}):, \tag{27.1}$$

where :: denotes Wick ordering. The constant e is given by

$$e := \sqrt{4\pi\alpha},$$

where $\alpha \approx 1/137$ is a dimensionless constant known as the *fine structure constant*. As always, we are working in units where $\hbar = c = 1$.

This has the usual mathematical ambiguities. For example, H_I is not really defined as an operator on our Hilbert space $\mathcal{R}$. But as before, we will pretend that it is, and carry on with our perturbative calculations based on the Dyson series expansion.

27.3. Perturbative calculations

We will now briefly see how to do perturbative calculations in this theory, without going into details (which would require much more time than we have left in this course). Consider the following scattering example. We have two incoming electrons with momenta $\mathbf{p}_1, \mathbf{p}_2$ and spins s_1, s_2. We want to compute the probability amplitude of getting two outgoing electrons with momenta $\mathbf{p}_3, \mathbf{p}_4$ and spins s_3, s_4.

As always, the incoming and outgoing states can be written using the creation operators. Specifically, we can write the incoming state as

$$a^\dagger(\mathbf{p}_1, s_1) a^\dagger(\mathbf{p}_2, s_2)|0\rangle$$

and the outgoing state as

$$a^\dagger(\mathbf{p}_3, s_3) a^\dagger(\mathbf{p}_4, s_4)|0\rangle.$$

The scattering amplitude

$$\langle 0|a(\mathbf{p}_3, s_3) a(\mathbf{p}_4, s_4) S a^\dagger(\mathbf{p}_1, s_1) a^\dagger(\mathbf{p}_2, s_2)|0\rangle$$

is defined as usual, where S is the scattering operator.

We can do the usual Dyson expansion. Suppose that $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4$ are distinct, so that there is no zeroth order term. For the higher terms, we need to work with a Hamiltonian density. The formula (27.1) suggests that the Hamiltonian density should be

$$\mathcal{H}(x) = e \sum_{\mu=0}^{3} :\overline{\psi}(x)\gamma^\mu \psi(x) A_\mu(x):.$$

Indeed, it can be easily verified that $\mathcal{H}$ satisfies the conditions required for Hamiltonian densities discussed in Lecture 16. Since there is no coupling

constant, the nth term in the Dyson expansion can be written as

$$\frac{(-i)^n}{n!}\int\dots\int dx_1^4\dots dx_n^4\langle 0|a(\mathbf{p}_3,s_3)a(\mathbf{p}_4,s_4)\\ \mathcal{T}\mathcal{H}(x_1)\mathcal{H}(x_2)\dots\mathcal{H}(x_n)a^\dagger(\mathbf{p}_1,s_1)a^\dagger(\mathbf{p}_2,s_2)|0\rangle, \quad (27.2)$$

where $\mathcal{H}$ is the Hamiltonian density defined above.

27.4. Feynman diagrams for quantum electrodynamics

Note that the integrand in (27.2) is a linear combination of many terms of the form

$$\langle 0|a(\mathbf{p}_3,s_3)a(\mathbf{p}_4,s_4)\mathcal{T}(:\overline{\psi}_{j_1}(x_1)\psi_{k_1}(x_1)A_{\mu_1}(x_1):\cdots\\ \cdots:\overline{\psi}_{j_n}(x_n)\psi_{k_n}(x_n)A_{\mu_n}(x_n):)a^\dagger(\mathbf{p}_1,s_1)a^\dagger(\mathbf{p}_2,s_2)|0\rangle. \quad (27.3)$$

As in φ^4 theory, we can apply Wick's theorem to (27.3). Using the fact that we have commutation/anticommutation relations for all of the creation and annihilation operators on $\mathcal{B},\mathcal{F},\mathcal{F}'$, and that ψ, A_μ are defined in terms of these creation and annihilation operators, we know (in principle) how to compute quantities such as

$$\langle 0|\mathcal{T}\overline{\psi}_j(x)\psi_k(y)|0\rangle,\ \langle 0|\mathcal{T}A_\mu(x)A_\nu(y)|0\rangle,\ \text{ etc.}$$

(We will see details of this in the next lecture.) Therefore, our goal now is to understand how to draw and analyze Feynman diagrams for quantum electrodynamics.

As before, each operator in (27.3) is represented by a node with a single edge hanging out. The slight difference is that the lines for the $A_\mu(x)$'s are wavy instead of straight. Again, as before, we simplify matters by fusing the three nodes corresponding to each x_i into a single node, which is simply labeled as x_i. Therefore, this node has two straight lines and one wavy line coming out of it. Such nodes are called internal nodes. The nodes corresponding to incoming and outgoing particles are called external nodes. Figure 27.1 shows all nodes and lines appearing in the first order of perturbation theory. Since there is an odd number of lines in the diagram shown in

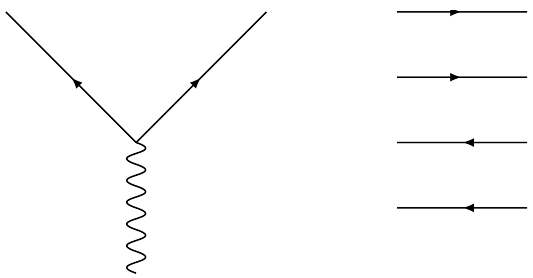

FIGURE 27.1. Vertices and edges for one internal node and four external nodes, occurring in the first term of the perturbation series

Figure 27.1, it is impossible to tie them without leaving at least one untied line. Therefore the first-order term is zero.

Let us next look at the second order term. The set of all nodes and lines for this term is shown in Figure 27.2. In this case it is in fact possible to

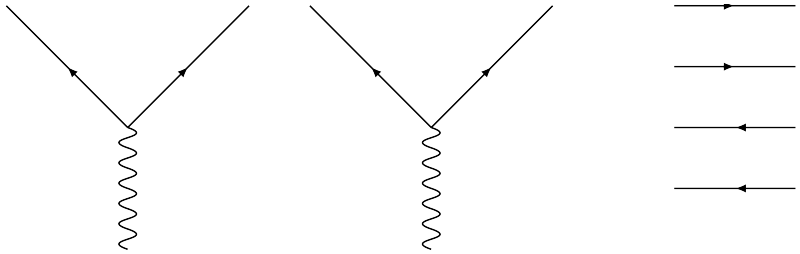

FIGURE 27.2. Vertices and edges for two internal nodes and four external nodes, occurring in the second term of the perturbation series

tie up all lines so that no line remains untied. The Wick ordering in the Hamiltonian density ensures that diagrams with self-loops at internal nodes do not contribute. A typical example of a contributing diagram is shown in Figure 27.3. Physically, it corresponds to two incoming electrons exchanging a photon and scattering into two outgoing electrons.

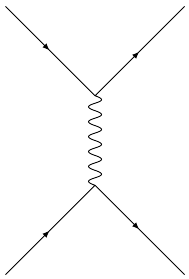

FIGURE 27.3. Example of a matching of 6 lines: 4 corresponding to electrons and 2 to photons

In fact, this is the only kind of diagram that contributes. The reason is as follows. From all the considerations mentioned above, the other kind of diagram that could have contributed are diagrams where the two incoming electron lines meet at one internal node, and the two outgoing electron lines meet at the other internal node. However, such diagrams give zero contribution, because

$$\langle 0|a(\mathbf{p},s)\psi_k(x)|0\rangle = 0$$

for any $\mathbf{p}$, s, k and x, as is apparent from the definition (26.2) of $\psi_k(x)$ and the facts that $\langle a(\mathbf{p},s)a(\mathbf{p}',s')|0\rangle = 0$ and $\langle 0|a(\mathbf{p},s)b^\dagger(\mathbf{p}',s')|0\rangle = 0$ for any $\mathbf{p}$, s, $\mathbf{p}'$ and s'.

LECTURE 28

Electron scattering

Date: 12/5/2018
Scribe: Cyrus Reza

28.1. The Dirac propagator

In this lecture, we will briefly outline a calculation in QED without going into the details of all the steps. Recall that we are trying to compute the second order term in the Dyson expansion for the scattering amplitude

$$\langle 0|a(\mathbf{p}_3, s_3)a(\mathbf{p}_4, s_4)Sa^\dagger(\mathbf{p}_1, s_1)a^\dagger(\mathbf{p}_2, s_2)|0\rangle \tag{28.1}$$

of getting two outgoing electrons with momenta $\mathbf{p}_3$, $\mathbf{p}_4$ and spins s_3, s_4, from two incoming electrons with momenta $\mathbf{p}_1$, $\mathbf{p}_2$ and spins s_1, s_2. We have argued that the only type of Feynman diagram that contributes to this calculation are diagrams of the type shown in Figure 27.3.

To compute the contribution of any internal line in QED, we need to calculate two kinds of *propagators* in QED, akin to the Feynman propagator appearing in φ^4 scalar field theory. Recall that the Feynman propagator for a particle of mass m is given by

$$\Delta_{F,m}(x) = \lim_{\varepsilon\to 0}\int \frac{d^4p}{(2\pi)^4}\frac{e^{-i(x,p)}}{-p^2+m^2-i\varepsilon}.$$

(We were writing it earlier as Δ_F, but we will need to be explicit about m here.) The *Dirac propagator* is a 4×4 array of Schwartz distributions, defined as

$$\Delta_D(x) = \sum_{\mu=0}^{3}(i\partial_\mu\Delta_{F,m}(x))\gamma^\mu + m\Delta_{F,m}(x)I,$$

where γ^μ are the Dirac matrices defined in the previous lecture, and I is the 4×4 identity matrix. The Dirac propagator appears in QED calculations in the following way. For any $1\le j,k\le 4$,

$$\langle 0|\mathcal{T}\psi_j(x_1)\overline{\psi}_k(x_2)|0\rangle = -\langle 0|\mathcal{T}\overline{\psi}_j(x_1)\psi_k(x_2)|0\rangle = -i(\Delta_D(x_1-x_2))_{jk}.$$

This kind of internal line does not, however, appear in Figure 27.3, but do appear in higher orders of perturbation theory.

28.2. The photon propagator

The photon propagator $\Delta_{\mu\nu}$, for $0 \leq \mu, \nu \leq 3$, is defined as

$$\Delta_{\mu\nu}(x) = -\eta_{\mu\nu}\Delta_{F,0}(x),$$

where $\Delta_{F,0}$ is the Feynman propagator with $m = 0$, and as usual,

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

The photon propagator appears in QED calculations in the following way. For any $0 \leq \mu, \nu \leq 3$,

$$\langle 0|\mathcal{T}A_\mu(x_1)A_\nu(x_2)|0\rangle = -i\Delta_{\mu\nu}(x_1 - x_2).$$

28.3. Second order term in electron scattering

The Dirac propagator and the photon propagator allow us to calculate the contributions of the internal lines in Feynman diagrams for QED. Contributions from external lines, which are of the form $\langle 0|\psi_k(x)a^\dagger(\mathbf{p}, s)|0\rangle$, etc., can be calculated using commutation relations for the creation and annihilation operators. Without going into details, let us just mention that the end result of such calculations show that the second order term in the Dyson expansion of (28.1) is

$$\sum_{\mu,\nu=0}^{3} \frac{im^2e^2}{\sqrt{16\omega_{\mathbf{p}_1}\omega_{\mathbf{p}_2}\omega_{\mathbf{p}_3}\omega_{\mathbf{p}_4}}}\overline{u}(\mathbf{p}_3, s_3)\gamma^\mu u(\mathbf{p}_1, s_1)\frac{\eta_{\mu\nu}}{(p_3 - p_1)^2}\overline{u}(\mathbf{p}_4, s_4)\gamma^\nu u(\mathbf{p}_2, s_2)$$
$$+ \text{similar terms},$$

where $\overline{u}(\mathbf{p}, s)$ is the Dirac adjoint of the vector $u(\mathbf{p}, s)$ defined in (26.1), and the "similar terms" are obtained by exchanging $\mathbf{p}_3$ with $\mathbf{p}_4$, and $\mathbf{p}_1$ with $\mathbf{p}_2$.

28.4. The non-relativistic limit

Let us now consider the non-relativistic limit of the scattering amplitude computed above, that is, where $|\mathbf{p}_i| \ll 1$ for each i. In this situation, we have $\omega_{\mathbf{p}_i} \approx m$ for each i. We also have

$$(p_3 - p_1)^2 = (p_3^0 - p_1^0)^2 - (\mathbf{p}_3 - \mathbf{p}_2)^2,$$

and

$$p_i^0 = \sqrt{m^2 + \mathbf{p}_i^2} \approx m + \frac{\mathbf{p}_i^2}{m},$$

which implies that

$$(p_3 - p_1)^2 = -(\mathbf{p}_3 - \mathbf{p}_1)^2(1 + o(1)).$$

Lastly,

$$\overline{u}(\mathbf{p}, s)\gamma^\mu u(\mathbf{p}', s') \approx \overline{u}(\mathbf{0}, s)\gamma^\mu u(\mathbf{0}, s'),$$

and the right side, when summed over μ, gives $2\delta_{s,s'}$. Therefore in the non-relativistic limit, the second order term in the scattering amplitude is approximately equal to

$$\frac{-ie^2}{(\mathbf{p}_3-\mathbf{p}_1)^2}\delta_{s_1,s_3}\delta_{s_2,s_4}(2\pi)^4\delta^{(4)}(p_3+p_4-p_1-p_2)+\text{similar terms.}$$

The function $V:\mathbb{R}^3\to\mathbb{R}$ whose Fourier transform is

$$\widehat{V}(p)=\frac{1}{\mathbf{p}^2}$$

is the function

$$V(x)=\frac{1}{4\pi|\mathbf{x}|}.$$

This indicates that if we assumed a non-relativistic quantum mechanical model for electrons and used the Coulomb potential to model the repulsion between electrons, we would have ended up with the same scattering amplitude. Thus, the Coulomb potential arises naturally in the non-relativistic limit of QED.

28.5. Anomalous magnetic moment of the electron

The anomalous magnetic moment of the electron is a dimensionless quantity, predicted to be equal to 1 by non-relativistic quantum mechanics ($\hbar=c=1$). The experimental value deviates slightly from this prediction. Currently, the experimentally determined value, up to ten places of decimal, is the following:

$$1.001\,159\,652\,180\,73(\pm 28).$$

It is one of the great successes of QED that it gives a much more accurate prediction for this quantity. The theoretical value, when calculated using second order perturbation theory, already gives

$$1.001\,161\,4.$$

However, when one goes to fifth order of perturbation theory, one gets

$$1.001\,159\,652\,181\,643(\pm 764),$$

which is a truly astonishing match with experiment, unparalleled in all of science.

LECTURE 29

The Wightman axioms

Date: 12/7/2018
Scribe: Chenyang Zhong

29.1. Interacting quantum fields

Until now, we have only worked with free fields, and scattering amplitudes computed using free fields. There are, however, important objects known as interacting quantum fields that are more complicated than free fields. Let us now see the simplest example of an interacting quantum field. Recall φ^4 theory. We have the Hilbert space $\mathcal{H} = L^2(X_m, d\lambda_m)$ and the scalar free field φ. In this lecture, we will denote the free field by φ^{free}, to distinguish it from the interacting field φ that we are going to define below.

Recall the Hamiltonian $H = H_0 + gH_I$, where H_0 is the Hamiltonian of free evolution, and

$$H_I = \frac{1}{4!}\int d^3\mathbf{x} {:}\varphi^{free}(0,\mathbf{x})^4{:}.$$

The *interacting field* φ defined by the Hamiltonian H is

$$\varphi(t,\mathbf{x}) := e^{itH}\varphi^{free}(0,\mathbf{x})e^{-itH}.$$

In particular, $\varphi(0,\mathbf{x}) = \varphi^{free}(0,\mathbf{x})$. In other words, the interacting field starts off as φ^{free} at time zero, but then evolves according to the Hamiltonian H instead of H_0 (in the Heisenberg picture of evolution of operators). In contrast, the free field evolves as

$$\varphi^{free}(t,\mathbf{x}) = e^{itH_0}\varphi^{free}(0,\mathbf{x})e^{-itH_0},$$

a fact that we have noted before.

Note that the definition of φ does not make mathematical sense, since H is not a mathematically well-defined operator on $\mathcal{H}$. The rigorous definition of interacting quantum fields is one of the major open questions in mathematical physics. We will discuss one approach to the solution of this question below.

29.2. Green's functions

Suppose that H is indeed a well-define operator (possibly on a space that is different than $\mathcal{H}$), and that it has a unique ground state (i.e. an eigenstate with minimum eigenvalue) Ω. This is called the vacuum state of the theory.

Note that $\Omega \neq |0\rangle$. The Green's functions (or Schwinger functions) for the theory are defined, for each n, as

$$G(x_1, \ldots, x_n) = \langle \Omega | \mathcal{T} \varphi(x_1) \cdots \varphi(x_n) | \Omega \rangle \tag{29.1}$$

where $x_1, \ldots, x_n \in \mathbb{R}^{1,3}$.

Typically, G is expected to be a distribution rather than a function. It turns out that the Green's functions contain much more information than scattering amplitudes. In particular, scattering amplitudes can be recovered from the Green's functions using the LSZ (Lehmann–Symanzik–Zimmermann) formula. The Green's functions for an interacting theory are computed using the free field and the Gell-Mann–Low theorem.

29.3. The Wightman axioms

The Wightman axioms give a framework for rigorously defining interacting quantum field theories. The axioms are designed to ensure that there are no physical inconsistencies. Let us only consider the axioms for bosonic scalar fields for now.

Axiom 1. There exists a Hilbert space $\mathcal{H}$ whose elements represent the possible states of the quantum system under consideration, and a unitary representation U of $\mathcal{P}$ in $\mathcal{H}$. (We do not need "projective" because we are dealing with scalar fields.) There is one further assumption about U, and it needs a bit of preparation to state. Note that $(U(a,1))_{a \in \mathbb{R}^{1,3}}$ is a commuting family of unitary operators. Therefore by the spectral theorem, we have that

$$U(a,1) = \int_{\mathbb{R}^{1,3}} dE(p) e^{i(a,p)},$$

where E is a projection-valued measure on $\mathbb{R}^{1,3}$. Let

$$L := \{p \in \mathbb{R}^{1,3} : p^0 \geq |\mathbf{p}|\}$$

be the "forward light cone". The additional assumption on U is that

$$E(A) = 0 \quad \text{whenever} \quad A \cap L = \emptyset.$$

This assumption is required to ensure that there are no particles with "negative masses" in the theory. There is a simple way to check (29.3). The condition is satisfied if and only if for any $\xi, \eta \in \mathcal{H}$ and any $f \in \mathscr{S}(\mathbb{R}^{1,3})$, such that the support of $\widehat{f}$ does not intersect L, we have

$$\int d^4a f(a)(\xi, U(a,1)\eta) = 0.$$

Axiom 2. There exists a dense subspace $\mathcal{D}$ of $\mathcal{H}$ and a linear map

$$\varphi : \mathscr{S}(\mathbb{R}^4) \to \{\text{unbounded operators on } \mathcal{H}\}$$

such that for $f \in \mathscr{S}(\mathbb{R}^4)\}$, we have that $\varphi(f)$ is defined on $\mathcal{D}$ and $\varphi(f)(\mathcal{D}) \subset \mathcal{D}$. Moreover, we assume that for any $\xi, \eta \in \mathcal{D}$,

$$(\varphi(f^*)\xi, \eta) = (\xi, \varphi(f)\eta).$$

In particular if f is real-valued, then $\varphi(f)$ is a symmetric operator on $\mathcal{D}$.

Axiom 3. For any $\xi, \eta \in \mathcal{D}$, $f \mapsto (\xi, \varphi(f)\eta)$ is a tempered distribution.

Axiom 4. For any $(a, A) \in \mathcal{P}$,

$$U(a, A)\varphi(f)U(a, A)^{-1} = \varphi(V(a, A)f),$$

where

$$V(a, A)f(x) = f(A^{-1}(x - a)).$$

This is the axiom of Lorentz invariance, designed to ensure compatibility with special relativity.

Axiom 5. If the supports of $f, g \in \mathscr{S}(\mathbb{R}^4)$ are "spacelike separated", then

$$[\phi(f), \phi(g)] = 0.$$

(Two points $x, y \in \mathbb{R}^{1,3}$ are called spacelike separated if $(x - y)^2 < 0$. Two sets are spacelike separated if any point in one spacelike separated from any point in the other.) This is the locality axiom, designed to ensure compatibility with the stipulation of special relativity that a signal cannot travel faster than the speed of light.

Axiom 6. There exists $\Omega \in \mathcal{D}$ such that $U(a, 1)\Omega = \Omega$ for any a, and it is the unique element satisfying this condition. (One can actually deduce from the axioms that $U(a, A)\Omega = \Omega$ for any $(a, A) \in \mathcal{P}$.) This is the axiom that ensures the existence and uniqueness of the vacuum state.

Axiom 7. The closed linear span of elements like $\varphi(f_1) \cdots \varphi(f_n)\Omega$ for n and $f_1, \ldots, f_n$ arbitrary, is the Hilbert space $\mathcal{H}$. This axiom is designed to ensure that the field φ is "rich enough" to generate all possible states.

It can be verified that the scalar free field satisfies all of the above axioms. A quantum field theory φ is called interacting if it is not the free field. The Wightman axioms have analogs in $\mathbb{R}^{1,1}$ and $\mathbb{R}^{1,2}$, where interacting fields satisfying these axioms have been constructed. Unfortunately, no one has been able to construct an interacting quantum field theory in $\mathbb{R}^{1,3}$ that satisfies the Wightman axioms.

29.4. The probabilistic approach

There is a probabilistic approach to the construction of interacting quantum field theories that satisfy the Wightman axioms. Suppose that we start

with the Green's function $G(x_1, \dots, x_n) = G((t_1, \mathbf{x}_1), \dots, (t_n \mathbf{x}_n))$ for a hypothetical quantum field theory, and analytically analytically continue it in time to get

$$S(x_1, \dots, x_n) := G((it_1, \mathbf{x}_1), \dots, (it_n, \mathbf{x}_n)).$$

From the path integral heuristic (which we have not discussed), it is expected that S is the n-point correlation function of a random field:

$$S(x_1, \dots, x_n) = \mathbf{E}(\psi(x_1) \cdots \psi(x_n)),$$

where ψ is a random field on $\mathbb{R}^{1,3}$. Usually the field ψ is so rough that S is still a distribution rather than a function. The correct random field can be identified using the Lagrangian of the theory (again, which we have not discussed).

Once the guess about the random field has been made, one can try to verify that it satisfies a set of axioms known as the Osterwalder–Schrader axioms. If these axioms are satisfied by ψ, then there exists a "reconstruction theorem" that reconstructs the quantum field whose Green's function G is the analytic continuation of S in time. Here reconstructing the quantum field means constructing a quantum field theory satisfying the Wightman axioms.

Using this approach, interacting scalar φ^4 quantum field theories have been constructed on $\mathbb{R}^{1,1}$ and $\mathbb{R}^{1,2}$, but not yet on $\mathbb{R}^{1,3}$.

Open problem 29.1. Construct an interacting scalar quantum field theory on $\mathbb{R}^{1,3}$ that satisfies the Wightman axioms.