PHYS 526 (Quantum Field Theory I) Notes

Rio Weil

This document was typeset on October 19, 2022

Introduction:

This is a set of lecture notes taken from UBC's PHYS 526 (Graduate Quantum Field Theory I) course, taught by Dr. Gordon Semenoff. The course covers many-particle systems, second quantization, degenerate Fermi and Bose gases, The action principle and Noether's theorem, Non-relativistic space-time symmetries, relativistic field theories, real scalar quantum field theory, Emergent relativistic symemtry, Dirac field theory, photons, functional methods, and perturbative quantum electrodynamics. If any errors are found in the notes, feel free to email me at ryoheiweil@phas.ubc.ca.

Contents

1 Motivation and Many Particles

1.1 Why QFT?

1. Natural way to study QM systems with large number of DOFs

2. To reconcile special relativity and quantum mechanics ("QM + SR = QFT"). No way to do "regular" QM in a relativistic setting. Mainly because $E = mc^2$, so energy can convert into mass; e.g. highly energetic collisions in the LHC which produce a large number of particles. You need a framework which can account for a large type and an arbitrary number of quantum-mechanical particles.

3. If you take point quantum mechanics and replace the NRQM Hamiltonian (with non-relativistic momentum) with the relativistic version of $H = \sqrt{p^2c^2 + m^2c^4} = mc^2 + \frac{p^2}{2m} + \cdots$, you find that the particle disperses (much like in the non-relativistic case) but it spreads in such a way that it violates causality; i.e. it can disperse outside of the light cone. There is no way to repair this in single-particle point quantum mechanics. QFT fixes this beautifully. It introduces an antiparticle, and says that the acausal process is actually a superposition of two processes, one with the particle and one with the antiparticle, and one "tune" the superposition so there is a destructive interference of the acausal behavior.

1.2 A one-particle QM system

Let's review a one-particle system. It is described by a wavefunction $\psi(\mathbf{x}, t)$, which satisfies the Schrodinger equation:

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{x}, t) = \left(-\frac{\hbar^2 \mathbf{\nabla}^2}{2m} + V(\mathbf{x}) \right) \psi(\mathbf{x}, t). \tag{1.1}$$

In the above equation, $H = -\frac{\hbar^2 \mathbf{\nabla}^2}{2m} + V(\mathbf{x})$ is the Hamiltonian (the so called "energy"), where the first term is the kinetic energy and the second term is the position-dependent potential energy (e.g. due to gravitational interaction, electronic interactions). Note we assume that the potential is velocity-independent to simplify things. We take the momentum to be:

$$\mathbf{p} := -i\hbar \mathbf{\nabla} \tag{1.2}$$

so the kinetic energy is of course:

$$\frac{\mathbf{p}}{2m} = -\frac{\hbar^2 \mathbf{\nabla}^2}{2m}. \tag{1.3}$$

The nabla operator is defined as $\mathbf{\nabla} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)$, and so $\mathbf{\nabla}^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$. Note that the Schrodinger equation is linear (and its validity can be confirmed in experiment, though we take it as an axiom in NRQM). To the wavefunction we can associate a probability amplitude:

$$|\psi(\mathbf{x}, t)|^2 d^3 x = \text{ Probability of finding particle in volume } d^3 x \text{ at position } \mathbf{x}, \text{ time } t. \tag{1.4}$$

And since we must find the particle somewhere, we have the normalization condition:

$$\int |\psi(\mathbf{x}, t)|^2 d^3 x = 1. \tag{1.5}$$

We have not yet specified where the particle is allowed to be. If the particle is confined to some region (e.g. a box) then we require the enforcement of boundary conditions on the wavefunction (e.g. $\psi(x, 0) = \psi(x, L) = 0$ for a infinite square well). For our purposes, we will take $\mathbf{x} \in \mathbb{R}^3$ (no confinement), with the boundary condition specified by the normalization condition (but sometimes we even relax this, e.g. with plane waves, where we might specify the BC as the existence of the Fourier transform).

1.3 A many-particle QM system

We now move to a many-particle quantum mechanical system. How does the Schrodinger equation look in this case? For an identical N-particle system, a natural generalization is:

$$i\hbar\frac{\partial}{\partial t}\psi(\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_N,t) = \left(\sum_{i=1}^{N}\frac{-\hbar^2\boldsymbol{\nabla}_i^2}{2m} + V(\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_N)\right)\psi(\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_N,t). \tag{1.6}$$

But why is this a natural generalization? Suppose Alice and Bob have a particle each, and are studying the particles in two far-apart labs. They both analyze their experiment using a one-particle Schrodinger equation (Alice should not have to take into account Bob's particle on the other side of the world, and vise versa! Physics should be local). Perhaps they are doing similar experiments, and start exchanging emails, and want to describe the system as a composite. The natural way to create a composite system would be to multiply the wavefunction of particle one by the wavefunction of particle two:

$$\psi(\mathbf{x}_1,\mathbf{x}_2,t) = \psi_1(\mathbf{x}_1,t)\psi_2(\mathbf{x}_2,t) \tag{1.7}$$

this follows naturally from the probabilistic interpretation of the wavefunctions (when we compose two probability distributions, we take their product, not their sum). We could then show that the product of the two wavefunctions satisfies the composite SE Eq. (1.6) if they satisfy their individual one-particle Schrodinger equations, i.e.:

$$i\hbar\frac{\partial}{\partial t}\psi_1(\mathbf{x}_1,t)\psi_2(\mathbf{x}_2,t) = \left(-\frac{-\hbar^2\boldsymbol{\nabla}^2}{2m} - \frac{\hbar^2\boldsymbol{\nabla}_2^2}{2m} + V(\mathbf{x}_1) + V(\mathbf{x}_2)\right)\psi_1(\mathbf{x}_1,t)\psi_2(\mathbf{x}_2,t). \tag{1.8}$$

However, if we introduce an interaction between the two particles (e.g. a Coloumb interaction), taking the composite wavefunction as the product no longer becomes valid; however Eq. (1.6) still holds.

We now return to the assumption that the particles are identical. Of course this means that $m_1 = m_2 = \ldots m_N$, but this has the more interesting implication that V is symmetric in its arguments, i.e.

$$V(\mathbf{x}_{P(1)},\mathbf{x}_{P(2)},\ldots,\mathbf{x}_{P(N)}) = V(\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_N) \tag{1.9}$$

for any permutation $(P(1),\ldots P(N))$ of $(1,\ldots N)$. There are $N!$ permutation of N objects. This is what it means for the particles to be identical, as they feel interactions in a way such that the interaction is left unchanged by swapping any of the particles.

2 Many Particles Continued, Second Quantization

2.1 Bosons and Fermions

Recall the many-particle Schrodinger Equation:

$$i\hbar\frac{\partial}{\partial t}\psi(\mathbf{x}_1,\ldots,\mathbf{x}_N,t) = \left(\sum_{i=1}^{N}\frac{-\hbar^2\boldsymbol{\nabla}_i^2}{2m} + V(\mathbf{x}_1,\ldots,\mathbf{x}_N)\right)\psi(\mathbf{x}_1,\ldots,\mathbf{x}_N,t). \tag{2.1}$$

This is the fundamental mathematical problem we are to solve when doing QM. There are depressingly few examples which are exactly solvable (almost none), such as single-particle potentials (where potentials such as the harmonic oscillator, hydrogen atom, infinite square well are exactly solvable). If we include two-particle potentials and $N \geq 3$, things are very difficult. There are a few low-dimensional examples solved with sophisticated techniques, and a couple solutions for $N = 2$, but that's about it. We've seen a few of these in prior QM courses. It may be depressing that we are talking about equations we can't solve, but we don't have to write down solutions; there are at least existence theorems for solutions (requiring boundary conditions/initial values; the initial value determines it uniquely and deterministically at later times).

We have gone out of our way to make the particles identical here; all particles have the same mass m, and the potential is a symmetric function of its arguments (it is unchanged under permutation of indices on the coordinates). This gives this equation a very high degree of symmetry[1]. This tells us that if we manage to find a solution, we obtain another solution by permuting the labels:

$$\psi(\mathbf{x}_1,\ldots,\mathbf{x}_N,t) \to \psi(\mathbf{x}_{P(1)},\ldots,\mathbf{x}_{P(N)},t) \tag{2.2}$$

so one solution gives us $N!$ solutions, and then by the principle of superposition we actually obtain an infinite number. But to be different quantum states, they should be linearly independent as vectors, i.e.:

$$c_1\psi + c_2\psi_P = 0 \tag{2.3}$$

can only be solved by $c_1 = c_2 = 0$. If they are linearly independent, then $\psi_P = -\frac{c_1}{c_2}\psi$ (or $e^{i\varphi}\psi$ if the states are normalized). So which is it? At this point, mathematics doesn't help us, but mother nature does come to the rescue and chooses one of these; nature says that they always have to be linearly dependent; so there is only one state[2]. After you find one solution, you have *not* found $N!$ solutions, but just the one. Given this, we can consider a particular interchange where we swap of the two labels. If we do it twice, we should come back to the same state:

$$\psi_P = -\frac{c_1}{c_2}\psi = \left(-\frac{c_1}{c_2}\right)^2\psi_P \tag{2.4}$$

so this tells us that $(-c_1/c_2)^2 = 1$, i.e. $-c_1/c_2$ is either 1 or -1. If we do an interchange of labels, we will have two cases:

$$\psi(\mathbf{x}_{P(1)},\ldots,\mathbf{x}_{P(N)},t) = \psi(\mathbf{x}_1,\ldots,\mathbf{x}_N,t) \tag{2.5}$$

where no change happens for any permutation; such particles are known as **bosons**. Or, we can have:

$$\psi(\mathbf{x}_{P(1)},\ldots,\mathbf{x}_{P(N)},t) = (-1)^{\deg(P)}\psi(\mathbf{x}_1,\ldots,\mathbf{x}_N,t) \tag{2.6}$$

where $\deg(P)$ is the number of neighbours necessary to interchange to put the labels back in order (this is easily seen to be defined mod 2). Such particles are known as **fermions**[3].

---

[1]Symmetry will be a central focus in future lectures.

[2]Perhaps the only time in life where nature picks the simplest path...

[3]These two are the only possibilities in three dimensions; in lower dimensions particles known as *anyons* are also possible (and useful for fault-tolerant topological quantum computing, as it turns out!), but these are outside the scope of the course. In relativistic physics, there are theorems that tell you particles are always bosons/fermions, but these theorems always have caveats; so in some sense this uniqueness of particle types comes down to what has been observed (and certainly this is the case for NRQM).

Note that this affects the counting of states quite severely. For indistinguishable particles, we have only one state (rather than a multiplicity of states) when we find a solution. Bosons are said to follow Bose-Einstein statistics, while Fermions follow Fermi-Dirac[4] statistics.

2.2 Particles with Spin

How do we account for the spin of particles? If we were just writing the SE for one particle with spin, we would write:

$$i\hbar\frac{\partial}{\partial t}\psi_\sigma(\mathbf{x},t) = \left(-\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2 + V(\mathbf{x})\right)\psi_\sigma(\mathbf{x},t) \tag{2.7}$$

where σ is a discrete index that runs over the possible spin polarizations of the particle:

$$\sigma = -J, -J+1, \ldots, J-1, J. \tag{2.8}$$

Of course the potential could have a dependence on the spin (e.g. spin-orbit coupling, spin-spin coupling):

$$\sum_{\tau=-J}^{\tau} V_\sigma^\tau(\mathbf{x})\psi_\tau(\mathbf{x},t). \tag{2.9}$$

Though we do not consider such interactions here, they are of immense importance in nuclear and AMO physics. If we have multiple particles with spin, our wavefunction can now be written as:

$$\psi_{\sigma_1,\ldots,\sigma_N}(\mathbf{x}_1,\ldots,\mathbf{x}_N,t). \tag{2.10}$$

If we want to symmetrize or anti-symmetrize, we permute both the position and the spin labels:

$$\psi_{\sigma_1,\ldots,\sigma_N}(\mathbf{x}_1,\ldots,\mathbf{x}_N,t) = \pm\psi_{\sigma_{P(1)},\ldots,\sigma_{P(N)}}(\mathbf{x}_{P(1)},\ldots,\mathbf{x}_{P(N)},t). \tag{2.11}$$

2.3 The Potential

Another comment is on the multi-particle potential energy function. If the particles do not interact with one another, we have:

$$V(\mathbf{x}_1,\ldots,\mathbf{x}_N) = \sum_{i=1}^{N} V(\mathbf{x}_i) \tag{2.12}$$

But we could also have the sum of two-body potentials:

$$V(\mathbf{x}_1,\ldots,\mathbf{x}_N) = \sum_{i=1}^{N} V(\mathbf{x}_i) + \sum_{i<j} V(\mathbf{x}_i,\mathbf{x}_j) \tag{2.13}$$

But if we study nuclear or condensed matter physics, we also have higher-body interactions:

$$V(\mathbf{x}_1,\ldots,\mathbf{x}_N) = \sum_{i=1}^{N} V(\mathbf{x}_i) + \sum_{i<j} V(\mathbf{x}_i,\mathbf{x}_j) + \sum_{i<j<k} V(\mathbf{x}_i,\mathbf{x}_j,\mathbf{x}_k) + \ldots. \tag{2.14}$$

Now we might think; how might we separate/determine these in a unique way? The experimentalist answer is to put particles in one, or two, or three (or more) at a time to determine the N-particle forces one at a time. For the purposes of our course, we will generally limit ourselves to studying up to two-body interactions. Three-body interactions and higher rarely do things for us (exception: nucleons in the nucleus).

---

[4]Named after their founders; though really it was Bose that wrote to Einstein first, and the F-D was really due to Pauli...

2.4 Second Quantization

We've written down a problem we will never solve; we have done this in order to rewrite the problem. This rewrite is required for a few reasons; first, we may have an open quantum system, so the number of particles is *not* fixed (we can get around it in the picture we have painted above, e.g. by using the average number of particles, but it isn't ideal). Another point to make is that for a finite number of particles and an infinite volume, we get zero density; we would prefer to describe things with finite (instead of zero) density. We could fix this in the current picture by putting space into a finite box, but again this is another band-aid we require. Perhaps the biggest reason for the rewrite is for mathematical elegance. We now leave physics behind for a few moments to construct a useful formalism.

We keep the spin, and write down an object $\psi_\sigma(\mathbf{x})$. Note that ψ here is *not* a wavefunction here. It is instead an operator. Operators operate on states, the most trivial of which is given by the empty vacuum $|0\rangle$. This is the state of our many-particle system with no particles in it. We will assume that $\psi_\sigma(\mathbf{x})$ has the property that:

$$\psi_\sigma(\mathbf{x})|0\rangle = 0. \tag{2.15}$$

And we will also assume that $|0\rangle$ is normalized:

$$\langle 0|0\rangle = 1. \tag{2.16}$$

Note that $|0\rangle$ is not the zero vector of our vector space (as the above normalization condition should make clear). We will also assume the existence (and uniqueness) of the dual space, which contains the bra $\langle 0|$ which satisfies:

$$\langle 0|\psi^{\dagger\sigma}(\mathbf{x}) = 0. \tag{2.17}$$

We then need only one more step; the commutation relations that these operators obey. We will assume that:

$$[\psi_\sigma(\mathbf{x}), \psi_\rho(\mathbf{y})] = 0, \quad \forall \mathbf{x}, \mathbf{y}, \sigma, \rho \tag{2.18}$$

Taking the Hermitian adjoint of the above, we obtain:

$$[\psi^{\dagger\sigma}(\mathbf{x}), \psi^{\dagger\rho}(\mathbf{y})] = 0, \quad \forall \mathbf{x}, \mathbf{y}, \sigma, \rho. \tag{2.19}$$

We need something that doesn't commute for these to be operators (rather than numbers); we have a relation reminiscent of the annhilation/creation operators of the quantum harmonic oscillator:

$$[\psi_\sigma(\mathbf{x}), \psi^{\dagger\rho}(\mathbf{y})] = \delta^\sigma_\rho \delta^3(\mathbf{x} - \mathbf{y}). \tag{2.20}$$

When we do relativistic physics, the contra and covariant coordinates (up/down labels) will become important; for now they are just labels without too much significance (just helps us to keep track). We can now consider a family of states:

$$|0\rangle, \psi^{\dagger\sigma}(\mathbf{x})|0\rangle, \psi^{\dagger\sigma}(\mathbf{x})\psi^{\dagger\rho}(\mathbf{y})|0\rangle, \dots \tag{2.21}$$

we can think of these as the basis vectors of our vector space, given the name the *Fock space*. A general vector in Fock space is given by a linear combination of the above basis vectors. Among these vectors, we can find the matrix elements of ψ_σ, ψ^σ and so on. We can now write down the density operator:

$$\rho(\mathbf{x}) = \psi^{\dagger\sigma}(\mathbf{x})\psi_\sigma(\mathbf{x}). \tag{2.22}$$

When we pair an covariant/contravariant quantity (up/down label), there is an implied sum (Einstein summation convention); we have omitted a $\sum_{\sigma=-J}^{J}$ in the above expression. Now, if we operate the density operator on the vacuum state, we have:

$$\rho(\mathbf{x})|0\rangle = \psi^{\dagger\sigma}(\mathbf{x})\psi_\sigma(\mathbf{x})|0\rangle = 0. \tag{2.23}$$

We can use the commutation algebra to look at an arbitrary basis vector and the action of $\rho(\mathbf{x})$ on it:

$$\rho(\mathbf{x})\psi^{\dagger\sigma_1}(\mathbf{x}_1)\dots\psi^{\dagger\sigma_N}(\mathbf{x}_N)|0\rangle = \sum_{i=1}^{N}\delta(\mathbf{x}-\mathbf{x}_i)\psi^{\dagger\sigma_1}(\mathbf{x}_1)\dots\psi^{\dagger\sigma_N}(\mathbf{x}_N)|0\rangle \tag{2.24}$$

In other words, the basis states are eigenstates of $\rho(\mathbf{x})$ with eigenvalues $\sum_{i=1}^{N}\delta(\mathbf{x}-\mathbf{x}_i)$.

Note that the basis states are not really physical quantum mechanical states; the particles are at fixed positions and at fixed spins. Another point to make is that since the ψ^σs commute with each other, it is a completely symmetric state, i.e. a state of bosons. It's not possible to say which bosons are sitting at which position, and has bose-einstein statistics built in. Nature has been very kind to us; if nature did not have such statistics, we would not be able to use such construction.

Next time, we will discuss what happens when the ψ_σs are fermions rather than bosons. We will also look for how we formulate the Schrodinger equation in this second quantization language.

A remark on the Fock space; it is a continuously infinite basis. Not clear that this is a separable Hilbert space, which is where we do QM in. One of the tenets is that the basis for such a space is countable (of which what we have is not). We could improve this by replacing the construction of $\psi^{\dagger\sigma}(\mathbf{x})$s with $\int d^x f_i(\mathbf{x})\psi^{\dagger\sigma}(\mathbf{x})$ where f_i are square integrable functions. But this is not actually discrete. You can use a Cantor diagonalization argument to show that there will be an uncountable number of states. To get a Hilbert space, you need to restrict yourself to states with a finite total number of particles. Then consider Cauchy sequences of such basis states, and consider all of the basis states plus limits of such Cauchy sequences, giving a Hilbert space. But this is a subtetly that we will not really consider for the remainder of the course.

3 Second Quantization Continued

3.1 Second Quantization in the Schrodinger Picture

Last time, we were in the middle of a mathematical construction. We established a Hilbert space[5], with creation/annhilation operators with commutation relations:

$$[\psi_\sigma(\mathbf{x}), \psi^{\dagger\rho}(\mathbf{y})] = \delta^\rho_\sigma \delta^3(\mathbf{x} - \mathbf{y}). \tag{3.1}$$

and

$$[\psi_\sigma(\mathbf{x}), \psi_\rho(\mathbf{y})] = [\psi^{\dagger\sigma}(\mathbf{x}), \psi^{\dagger\rho}(\mathbf{y})] = 0. \tag{3.2}$$

and we could obtain basis states by operating (mutually commuting)[6] creation operators on the vacuum state:

$$\psi^{\dagger\sigma_1}(\mathbf{x}_1) \dots \psi^{\dagger\sigma_N}(\mathbf{x}_N)|0\rangle. \tag{3.3}$$

where the vacuum state is normalized:

$$\langle 0|0\rangle = 1 \tag{3.4}$$

and is sent to zero by the anhilation operators:

$$\psi_\sigma(\mathbf{x})|0\rangle = 0, \langle 0|\psi^{\dagger\sigma}(\mathbf{x}) = 0. \tag{3.5}$$

Now, consider a superposition:

$$\int d^3x_1 \dots d^3x_N \psi_{\sigma_1\dots\sigma_N}(\mathbf{x}_1, \dots, \mathbf{x}_N, t)\psi^{\dagger\sigma_1}(\mathbf{x}_1) \dots \psi^{\dagger\sigma_N}(\mathbf{x}_N)|0\rangle. \tag{3.6}$$

where one can think of $\psi_{\sigma_1\dots\sigma_N}(\mathbf{x}_1, \dots, \mathbf{x}_N, t)$ as the coefficients of the sum (though of course the labels vary continuously, so we have integrals instead) - we will see shortly that this is the wavefunction of the system. Of course we also implicitly sum over the spin labels according to the Einstein summation convention. We will worry about normalization later on. Note that we still have not done anything physical here, so let us now do that; we want to set up our many-particle Schrodinger Equation in this second quantization language. Let us give the integral in Eq. (3.6) a name; let's call it $|\Psi(t)\rangle$. We want to find some sort of equation which tells us how this object evolves in time. This should be given by:

$$i\hbar \frac{\partial}{\partial t}|\Psi(t)\rangle = H|\Psi(t)\rangle. \tag{3.7}$$

such that the above equation is equivalent to the old Schrodinger Equation (1.6); but note that it is written much more economically. In order to do so, we need to specify the Hamiltonian. In order to do so, we use the fact that the potential term in the SE can be decomposed into one, two, three, and in general N body interaction terms. We obtain:

$$
\begin{aligned}
H = \int d^3x \frac{\hbar^2}{2m}\boldsymbol{\nabla}\psi^{\dagger\sigma}(\mathbf{x}) \cdot \boldsymbol{\nabla}\psi_\sigma(\mathbf{x}) &+ \int d^3x V(\mathbf{x})\psi^{\dagger\sigma}(\mathbf{x})\psi_\sigma(\mathbf{x}) \\
&+ \frac{1}{2!}\int d^3x d^3y V(\mathbf{x}, \mathbf{y})\psi^{\dagger\sigma}(\mathbf{x})\psi^{\dagger\rho}(\mathbf{y})\psi_\rho(\mathbf{y})\psi_\sigma(\mathbf{x}) \\
&+ \frac{1}{3!}\int d^3x d^3y d^3z V(\mathbf{x}, \mathbf{y}, \mathbf{z})\psi^{\dagger\sigma}(\mathbf{x})\psi^{\dagger\rho}(\mathbf{y})\psi^{\dagger\tau}(\mathbf{z})\psi_\tau(\mathbf{z})\psi_\rho(\mathbf{y})\psi_\sigma(\mathbf{x}) \\
&+ \dots
\end{aligned} \tag{3.8}
$$

Clearly if we did not have a decomposition of the potential into N-body terms, this decomposition would not work out. Note that in the above Hamiltonian, we have assumed that the potential interaction doesn't

---

[5]Small mathematical detail we will not worry about; quantum states belong to a projective space, rather than a true vector space.
[6]If the particles were somehow distinguishable, this entire construction would fail; this commutativity would not make sense.

care about the spin (but if it did, we would make the Vs into a matrix, which depends on the spin; see the textbook for a more general formula). The terms in the above sum get successively complex (three body interactions are already a nightmare[7]), but for most physical scenarios we only have up to two-body interactions.

We look back at Eq. (3.7); on the LHS the time derivative affects only the $\psi_{\sigma_1 \ldots \sigma_N}(\mathbf{x}_1, \ldots, \mathbf{x}_N, t)$. On the RHS we get a mixture of terms from H acting on $|\Psi(t)\rangle$. We could then collect terms to see how the various terms act on $\psi_{\sigma_1 \ldots \sigma_N}(\mathbf{x}_1, \ldots, \mathbf{x}_N, t)$; doing so, we would completely reproduce the old N-body Schrodinger equation in (1.6), with the only caveat that we have not specified the particle number. To this end, we consider the number operator:

$$\mathcal{N} = \int d^3x \, \psi^{\dagger \sigma}(\mathbf{x}) \psi_\sigma(\mathbf{x}) \tag{3.9}$$

Which when acted on an arbitrary state $|\Psi(t)\rangle$ counts the particle number. So we could supplement Eq. (3.7) with Eq. (3.9), and pairing this with knowledge of what $\psi^{\dagger \sigma}(\mathbf{x})$, $\psi_\sigma(\mathbf{x})$ are from the commutation relations, we have a completely equivalent formulation of quantum mechanics. Note that we have really done nothing here, just rewrote the same thing in a different language.

We have established an example of a non-relativistic quantum field theory. There is a further step we can take to write this down, however. Note that all of our work above was in the Schrodinger picture, where the states are time-dependent and the operators are time-independent. We will find it useful to recast this in the Heisenberg picture[8], where the states are time-independent and the operators are time-dependent. Let us begin this reconstruction now.

3.2 Review of the Heisenberg Picture

We can write the time-dependent quantum state that solves Eq. (3.7) as:

$$|\Psi(t)\rangle = e^{-\frac{i}{\hbar}Ht}|\psi(0)\rangle. \tag{3.10}$$

Where $|0\rangle$ is the initial state of the system at time zero. Now, the expectation value of an operator $\mathcal{O}$ in the Schrodinger picture can be written as:

$$\langle \mathcal{O} \rangle (t) = \langle \Psi(t)|\mathcal{O}|\Psi(t)\rangle. \tag{3.11}$$

Now, substituting in Eq. (3.10) to the above, we have:

$$\langle \mathcal{O} \rangle (t) = \langle \Psi(0)|e^{\frac{i}{\hbar}Ht}\mathcal{O}e^{-\frac{i}{\hbar}Ht}|\psi(0)\rangle \tag{3.12}$$

So now we can define the time dependent operator:

$$\mathcal{O}(t) = e^{\frac{i}{\hbar}Ht}\mathcal{O}e^{-\frac{i}{\hbar}Ht} \tag{3.13}$$

and hence the expectation value can be written in the Heisenberg picture as:

$$\langle \mathcal{O} \rangle (t) = \langle \Psi(0)|\mathcal{O}(t)|\psi(0)\rangle. \tag{3.14}$$

And the time-dependence of the operators are described by the Heisenberg equation of motion:

$$\frac{\partial}{\partial t}\mathcal{O}(t) = \frac{i}{\hbar}[H, \mathcal{O}(t)]. \tag{3.15}$$

For most QM problems, this is a much harder picture to solve problems in; however this is the way we will proceed, as things will look much more like a QFT in this formalism.

---

[7]They are of interest in nuclear physics, but we won't generally concern ourselves with them here
[8]Though note that this rewrite would not be doing anything physically

3.3 Second Quantization in the Heisenberg Picture

In the Heisenberg picture, we can study the time evolution of the annhilation/creation operators:

$$i\hbar \frac{\partial}{\partial t} \psi_\sigma(\mathbf{x}, t) = [\psi_\sigma(\mathbf{x}, t), H].$$ (3.16)

Note that H here is time independent, but ψ_σ has acquired a time-dependence in the Heisenberg picture. Now taking the commutation relations for ψ_σ from our previous construction, we find:

$$[\psi_{\sigma_1}(\mathbf{x}_1, t), \psi_{\sigma_2}(\mathbf{x}_2, t)] = 0$$ (3.17)

in the case where the time ts are the same. It is the same story with the $\psi^{\dagger\sigma}$s:

$$[\psi^{\dagger\sigma_1}(\mathbf{x}_1, t), \psi^{\dagger\sigma_2}(\mathbf{x}_2, t)] = 0$$ (3.18)

and the final commutation relation:

$$[\psi_{\sigma_1}(\mathbf{x}_1, t), \psi^{\dagger\sigma_2}(\mathbf{x}_2, t)] = \delta^{\sigma_2}_{\sigma_1} \delta^3(\mathbf{x}_1 - \mathbf{x}_2).$$ (3.19)

In order to remind us that the times should be the same when we look at the above operators, we call these the *equal-time commutation relations*. They give us enough algebra to figure out what (3.16). So let us do just that.

$$i\hbar \frac{\partial}{\partial t} \psi_\sigma(\mathbf{x}, t) = -\frac{\hbar^2 \boldsymbol{\nabla}^2}{2m} \psi_\sigma(\mathbf{x}, t) + V(\mathbf{x}) \psi_\sigma(\mathbf{x}, t)$$
$$+ \int d^3 y V(\mathbf{x}, \mathbf{y}) \psi^{\dagger\rho}(\mathbf{y}) \psi_\rho(\mathbf{y}) \psi_\sigma(\mathbf{x}, t) + \dots$$ (3.20)

This is a lot simpler, and it *looks* like $\psi_\sigma(\mathbf{x}, t)$ satisfies a Schrodinger equation up to the second term (even though it is an operator and not a wavefunction). However then it becomes nonlinear as we add more terms (recall that the SE is linear). It does appear to satisfy the propogation of waves, and we will call it a *field equation*. So, in the Heisenberg picture, we obtain an equivalent formulation of quantum mechanics by imposing the equal-time commutation relations and the field equation, as well as the number operator constraint:

$$\mathcal{N}|\Psi\rangle = N|\Psi\rangle.$$ (3.21)

So this is an example of a full-blown quantum field theory, which we could turn around and use field-theoretic methods to attack. It is very easy to generalize this to the case where we have an infinite number of particles. It is also easy to generalize this to a system where the number of particles is not fixed (where we get rid of the constraint imposed by Eq. (3.21)).

3.4 From Bosons to Fermions

Everything we said here concerns bosons (from the construction of the operators to the Fock space states). For fermions, we want totally antisymmetric states. In order to do this, we simply replace the commutation relations of the annhilation/creation operators with anticommutation relations. By this one simple change, we obtain the fermionic field theory rather than the bosonic field theory.

Also note; if we have multiple types of particles, we have multiple field equations, with cross terms in the equations of motion; but we will not be too concerned about this scenario here.

A fun note: the number of fermions is either even or odd. There is no dynamical process which can change the parity (fermion superselection rule; important for (e.g.) condensed matter in the study of topological insulators). The flip of the sign of the state should not change anything about the universe (irrelevancy of the global phase).

Next time, we will try to solve a simple many-particle system in the Heisenberg picture. We will see that it is not as abstract as it currently looks!

4   Weakly Interacting Particles

4.1   Review of QFT construction

To briefly recap, we have taken the N-particle SE:

$$i\hbar\frac{\partial}{\partial t}\psi_{\sigma_1...\sigma_N}(\mathbf{x}_1,\ldots,\mathbf{x}_N,t) = \left(\sum_{i=1}^{N}-\frac{\hbar^2\boldsymbol{\nabla}_i^2}{2m} + \sum_{i<j}\lambda\delta(\mathbf{x}_i-\mathbf{x}_j)\right)\psi_{\sigma_1...\sigma_N}(\mathbf{x}_1,\ldots,\mathbf{x}_N,t) \tag{4.1}$$

(here with a short-range repulsive interaction) with the normalization condition:

$$\int d^3x_1\ldots d^3x_N\psi^{\dagger\sigma_1...\sigma_N}(\mathbf{x}_1,\ldots\mathbf{x}_N,t)\psi_{\sigma_1...\sigma_N}(\mathbf{x}_1,\ldots\mathbf{x}_N,t) = 1. \tag{4.2}$$

and we have shown that it is completely equivalent to the field equation:

$$\left(i\hbar\frac{\partial}{\partial t} + \frac{\hbar^2\boldsymbol{\nabla}^2}{2m}\right)\psi_\sigma(\mathbf{x},t) = \lambda\psi^{\dagger\rho}(\mathbf{x},t)\psi_\rho(\mathbf{x},t)\psi_\sigma(\mathbf{x},t) \tag{4.3}$$

where the ψ_σ are field operators with equal-time commutation relations:

$$[\psi_\sigma(\mathbf{x},t),\psi^{\dagger\rho}(\mathbf{y},t)] = \delta_\sigma^\rho\delta^3(\mathbf{x}-\mathbf{y}). \tag{4.4}$$

(the other commutation relations are zero). Actually, the two are almost equivalent; we also have to put in the number operator:

$$\mathcal{N} = \int d^3\psi^{\dagger\sigma}(\mathbf{x},t)\psi_\sigma(\mathbf{x},t) \tag{4.5}$$

which specifies the number of particles in the system; note that $\mathcal{N}$ commutes with H and so is time-independent (and we can take the t in the above expression to be whatever we like; useful when we want to apply the equal-time commutation relations). We will in the near future find a more clever way to show that this is time-independent.

The field equation and the operators already give us a QFT; the number operator is auxilary and only necessary for a complete equivalence to the old formulation. We can also bring the information about the Hamiltonian along with us (useful as the eigenvalues of the Hamiltonian provide valuable information):

$$H = \int d^3x\frac{\hbar^2}{2m}\boldsymbol{\nabla}\psi^{\dagger\sigma}(\mathbf{x},t)\cdot\boldsymbol{\nabla}\psi_\sigma(\mathbf{x},t) + \frac{\lambda}{2}\psi^{\dagger\sigma}(\mathbf{x},t)\psi^{\dagger\rho}(\mathbf{x},t)\psi_\rho(\mathbf{x},t)\psi_\sigma(\mathbf{x},t). \tag{4.6}$$

But again we will find a more sophisticated way to derive these later (from the field equations and commutation relations) using the fact that quantities are conserved.

Even for this very simple interaction potential, it is impossible to solve this analytically (for 1-D some work may be possible). So what do we do? Perhaps we can solve an approximation to it; we look at the limit in which the interaction is small/weak. We should have to define what "small" means, but let us avoid that for the moment.

In statistical mechanics, we never say the particles are never completley non-interacting, as the particles need to transfer energy in order to reach thermal equilibrium. We assume the interaction is there, but quite weak. But similar to that case, at a first pass we can just throw away the interaction and solve the non-interacting scenario.

4.2 Weakly (Non) Interacting Particles (Bosons)

In the limit of no interaction ($\lambda \to 0$) we have:

$$\left(i\hbar\frac{\partial}{\partial t} + \frac{\hbar^2 \nabla^2}{2m}\right)\psi_\sigma(\mathbf{x}, t) = 0. \tag{4.7}$$

Now we have a linear rather than a nonlinear PDE to solve. Not only this, but the one above is very easy to solve; there is no x or t dependence in the above, so we can solve it simply by a Fourier transform. Plugging in a plane wave, we have:

$$\left(i\hbar\frac{\partial}{\partial t} + \frac{\hbar^2 \nabla^2}{2m}\right)e^{i\mathbf{k}\cdot\mathbf{x} - \frac{i}{\hbar}\frac{\hbar k^2}{2m}t} = 0 \tag{4.8}$$

where the above equation is easily verified by the observations that $\frac{\partial}{\partial t}e^{i\omega t} = i\omega e^{i\omega t}$ and $\nabla e^{i\mathbf{k}\cdot\mathbf{x}} = i\mathbf{k}e^{i\mathbf{k}\cdot\mathbf{x}}$. However note we have really found a continuum of solutions, as the above works for any value of $\mathbf{k}$. At this point we need to convince ourselves that we have found the complete set of solutions, but of course the set of plane waves are complete so we have just that. A most general solution is written as a linear combination:

$$\psi_\sigma(\mathbf{x}, t) = \int \frac{d^3 k}{(2\pi)^{3/2}} e^{i\mathbf{k}\cdot\mathbf{x} - i\frac{\hbar k^2}{2m}t} a_\sigma(\mathbf{k}) \tag{4.9}$$

where the $a_\sigma(\mathbf{k})$ can be thought of the coefficients of the expansion. $\psi^{\dagger\sigma}(\mathbf{x}, t)$ can then easily be found to be:

$$\psi^{\dagger\sigma}(\mathbf{x}, t) = \int \frac{d^3 k}{(2\pi)^{3/2}} e^{-i\mathbf{k}\cdot\mathbf{x} + i\frac{\hbar k^2}{2m}t} a^{\dagger\sigma}(\mathbf{k}). \tag{4.10}$$

Using the commutation relations for the field operators, we can then find the commutation relations for the $a_\sigma / a^{\dagger\sigma}$ to be:

$$[a_\sigma(\mathbf{k}), a_\rho(\mathbf{l})] = [a^{\dagger\sigma}(\mathbf{k}, a^{\dagger\rho}(\mathbf{l})] = 0 \tag{4.11}$$

$$[a_\sigma(\mathbf{k}), a^{\dagger\rho}(\mathbf{l})] = \delta_\sigma^\rho \delta^3(\mathbf{k} - \mathbf{l}). \tag{4.12}$$

which is extremely similar in form to the commutation algebra of the ψs. We can then construct the vector space that the as act on. Letting $|0\rangle$ be the empty vacuum state, we have:

$$a_\sigma(\mathbf{k})|0\rangle = 0, \forall \mathbf{k}, \sigma \tag{4.13}$$

and the vector space has a basis composed of vectors of the form:

$$a^{\dagger\sigma_1}(\mathbf{k}_1) \dots a^{\dagger\sigma_N}(\mathbf{k}_N)|0\rangle. \tag{4.14}$$

The dual statement of the above is:

$$\langle 0|a^{\dagger\sigma}(\mathbf{k}) = 0, \forall \mathbf{k}, \sigma \tag{4.15}$$

We can now calculate matrix elements using the commutation algebra:

$$\langle 0|a_{\sigma_1}(\mathbf{k}_1) \dots a_{\sigma_m}(\mathbf{k}_m)a^{\dagger\rho_1}(\mathbf{l}_1) \dots a^{\dagger\rho_m}(\mathbf{l}_n)|0\rangle = \delta_{mn} \sum_P \delta(\mathbf{k} - \mathbf{l}_{P(1)})\delta_{\sigma_1}^{\rho_{\sigma(1)}} \dots \delta(\mathbf{k}_n - \mathbf{l}_{P(m)})\delta_{\sigma_n}^{\rho_{P(m)}} \tag{4.16}$$

which is messy, but really comes from the fact that the matrix element is symmetric in its arguments of $\mathbf{k}_i, \mathbf{l}_j$. Sometimes due to this symmetry we enforce a $\frac{1}{N!}$ to normalize for all permutations (which can be useful for some applications). This concludes the boson story, but what about fermions?

14

4.3 Weakly (Non) Interacting Particles (Fermions)

For bosons, we initially enforced symmetry of the wavefunction in the arguments. For fermions, we enforce antisymmetry instead. All commutators become anticommutators; and the structure of the above argument holds basically exactly the same with the commutation relations for ψ replaced with anticommutation relations. When we get to the as, we also replace the commutators with anticommutators.

When we compute the matrix elements, we get negative signs from the anticommutation relations, so:

$$\langle 0 | a_{\sigma_1}(\mathbf{k}_1) \ldots a_{\sigma_m}(\mathbf{k}_m) a^{\dagger \rho_1}(\mathbf{l}_1) \ldots a^{\dagger \rho_m}(\mathbf{l}_n) | 0 \rangle = \delta_{mn} \sum_P (-1)^{\deg(P)} \delta(\mathbf{k} - \mathbf{l}_{P(1)}) \delta_{\sigma_1}^{\rho_{\sigma(1)}} \ldots \delta(\mathbf{k}_n - \mathbf{l}_{P(m)}) \delta_{\sigma_n}^{\rho_{P(m)}} \tag{4.17}$$

There's not much of a difference so far; but we will find the many-particle states are profoundly different for bosons and fermions.

4.4 Understanding our solution to the theory

We now ask: in what sense have we "solved" this theory? To start, we can take our plane wave solution and plug it into the number and Hamiltonian operators. Doing so, we obtain:

$$\mathcal{N} = \int d^3 k \, a^{\dagger \sigma}(\mathbf{k}) a_\sigma(\mathbf{k}) \tag{4.18}$$

$$H = \int d^3 k \frac{\hbar^2 \mathbf{k}^2}{2m} a^{\dagger \sigma}(\mathbf{k}) a_\sigma(\mathbf{k}). \tag{4.19}$$

We find that this closely resembles the harmonic oscillator, and also that $\mathcal{N}$ and H are explicitly time-independent. Also, if we take our states (Eq. (4.14)), we see that they are indeed eigenstates of these operators (can be determined through the commutation algebra, or by interpreting these as harmonic oscillator eigenstates[9]):

$$\mathcal{N} a^{\dagger \sigma_1}(\mathbf{k}_1) \ldots a^{\dagger \sigma_N}(\mathbf{k}_N) | 0 \rangle = N a^{\dagger \sigma_1}(\mathbf{k}_1) \ldots a^{\dagger \sigma_N}(\mathbf{k}_N) | 0 \rangle \tag{4.20}$$

$$H a^{\dagger \sigma_1}(\mathbf{k}_1) \ldots a^{\dagger \sigma_N}(\mathbf{k}_N) | 0 \rangle = \left(\sum_{i=1}^N \frac{\hbar^2 \mathbf{k}_i^2}{2m} \right) \left(a^{\dagger \sigma_1}(\mathbf{k}_1) \ldots a^{\dagger \sigma_N}(\mathbf{k}_N) | 0 \rangle \right). \tag{4.21}$$

Note that this discussion applies equally as well to bosons and fermions.

4.5 Degenerate Fermi Gas - Vacuum State

We have found a complete solution in the non-interacting limit, from which we can draw out some information. If we think about a high-energy state where bosons/fermions can be treated the same, can we derive familiar results (e.g. high T limit for ideal gas/ideal gas law)? As a teaser for next time: we will start to study fermionic systems, which are slightly easier to understand. Consider a state with an infinite number of fermionic particles (as we want a finite density, if we have infinite space then we need infinite particles; in CM you may instead consider a finite number of particles in a box with boundary conditions). We will like to look at the low-energy states of such a system. If I had only one particle, its lowest energy state would be $\mathbf{k} = \mathbf{0}$, the state with a constant wavefunction. If I had two fermions, the first can have energy zero, but the second one cannot; this is because if we had two fermions in the same state, then $\{a^{\dagger \sigma}(\mathbf{k}) a^{\dagger \sigma}(\mathbf{k})\} = (2a^{\dagger \sigma}(\mathbf{k}))^2 = 0$ (although we could have two different spin states); so $a^{\dagger \sigma}(\mathbf{k}))^2 | 0 \rangle = 0$ is the zero vector and hence not a real quantum state. This is of course the famous *Pauli exclusion principle*.

---

[9]The quote that "space is just a bunch of harmonic oscillators" is salient here.

The next $\mathbf{k}$ above zero is slightly ill-defined in the infinite-space limit (as we have a continuum)[10] but let us imagine it. The N-particle ground state would be:

$$|\mathcal{O}\rangle = \prod_{\mathbf{k} < k_F, \sigma} \left(a^{\dagger\sigma}(\mathbf{k}) \right) |0\rangle. \tag{4.22}$$

Note that the above is really mathematical nonsense due to the continuity of $\mathbf{k}$. We call $|\mathcal{O}\rangle$ the vacuum (different from the empty vacuum $|0\rangle$). Note that the vacuum state we will stick with for the rest of the course, while the empty vacuum we will abandon when we get to relativistic field theory; it is not accessible in that limit. Let us define the vacuum state algebraically instead of the nonsense definition we have above (though we can heuristically understand the below algebraic constraints based on the nonsense equation we have above):

$$a^{\dagger\sigma}(\mathbf{k})|\mathcal{O}\rangle = 0 \text{ if } |\mathbf{k}| \leq k_F \tag{4.23}$$

$$a_\sigma(\mathbf{k})|\mathcal{O}\rangle = 0 \text{ if } |\mathbf{k}| > k_F. \tag{4.24}$$

$$\langle \mathcal{O}|\mathcal{O}\rangle = 0. \tag{4.25}$$

Note that the $\leq$ does not really matter in the above equation; the Fermi surface where $|\mathbf{k}| = k_F$ is a set of zero measure. k_F here is known as the Fermi wavenumber and from this we can construct $\hbar k_F$ the Fermi momentum, and $\epsilon_F = \frac{\hbar^2 k_F^2}{2m}$ the Fermi energy.

This construction is how we will deal with fermions next day!

---

[10]Here it might be nice to be a CM physicist instead...

5 Degenerate Fermi Gas

5.1 Review of the Weakly Interacting Fermion Construction

Last time, we started looking at the Fermi gas; it is what you get when you throw away the interactions between fermions (if you retain the interactions, you have something known as a Fermi liquid). Recall that we quantized the field:

$$\psi_\sigma(\mathbf{x}, t) = \int \frac{d^3k}{(2\pi)^{3/2}} e^{i\mathbf{k}\cdot\mathbf{x} - i\frac{\hbar \mathbf{k}^2}{2m} t} \alpha_\sigma(\mathbf{k}). \tag{5.1}$$

We recall that we showed how no two fermions could occupy the same state. So building up the ground state of an N-particle fermion state, we start with the lowest energy state and place one (actually two, accounting for spin) particle per energy level. Since $\epsilon \propto \mathbf{k}^2$, all states up to a spherical surface known as the *Fermi sphere* become filled.

We tried to construct this state (the vacuum), but it resulted in some mathematical ugliness (with an infinite continuous product of creation operators up to some $\mathbf{k}$... mathematical nonsense). But we got around this by talking about the state directly and postulating its properties. We say that:

$$\begin{aligned} \alpha_\sigma^\dagger(\mathbf{k})|\mathcal{O}\rangle &= 0 \quad |\mathbf{k}| \le k_F \\ \alpha_\sigma(\mathbf{k})|\mathcal{O}\rangle &= 0 \quad |\mathbf{k}| > k_F. \end{aligned} \tag{5.2}$$

though whether the states on the Fermi sphere are filled or empty will not be of particular importance to us.

5.2 Particles and Holes

Let us now rename some things:

$$\begin{aligned} \alpha_\sigma(\mathbf{k}) &= a_\sigma(\mathbf{k}) \quad \text{if } |\mathbf{k}| > k_F \\ \alpha_\sigma(\mathbf{k}) &= b_\sigma^\dagger(-\mathbf{k}) \quad \text{if } |\mathbf{k}| < k_F \end{aligned} \tag{5.3}$$

so it looks like we have two kinds of creation/annhilation operators. They are in a sense the same, but we label them like this to make the defining formulas for $|\mathcal{O}\rangle$ simpler, as now we can write them as:

$$\begin{aligned} a_\sigma(\mathbf{k})|\mathcal{O}\rangle &= 0 \quad |\mathbf{k}| > k_F \\ b_\sigma^\dagger(\mathbf{k})|\mathcal{O}\rangle &= 0 \quad |\mathbf{k}| < k_F. \end{aligned} \tag{5.4}$$

And we can write an excited state as:

$$(a^{\dagger \sigma_1}(\mathbf{k}_1) \dots b_{\rho_1}^\dagger(\mathbf{l}_1))|\mathcal{O}\rangle \tag{5.5}$$

where the as create fermions, and the bs create holes (annihilating what was already there). We can therefore write the quantized field operator as:

$$\psi_\sigma(\mathbf{x}, t) = \int_{|\mathbf{k}|>k_F} \frac{d^3k}{(2\pi)^{3/2}} e^{i\mathbf{k}\cdot\mathbf{x} - i\frac{\hbar\mathbf{k}^2}{2m}t} a_\sigma(\mathbf{k}) + \int_{|\mathbf{k}|<k_F} \frac{d^3k}{(2\pi)^{3/2}} e^{-i\mathbf{k}\cdot\mathbf{x} - i\frac{\hbar\mathbf{k}^2}{2m}t} b_\sigma^\dagger(\mathbf{k}) \tag{5.6}$$

where we have done a change of variables $\mathbf{k} \to -\mathbf{k}$ in the second integral. So, ψ has two parts; it either annhilates a particle, or creates a hole. The only funny part of the expression is that the phases have opposite signs, but the energy part does not. It looks like the holes somehow gives us back negative energy. There is something we can do to repair this; let us redefine the energy. As it is currently, we have $\epsilon = \frac{\hbar^2 \mathbf{k}^2}{2m}$. What if instead we add a constant[11]:

$$\frac{\hbar^2 \mathbf{k}^2}{2m} \to \frac{\hbar^2 \mathbf{k}^2}{2m} - \frac{\hbar^2 k_F^2}{2m} = \frac{\hbar^2 \mathbf{k}^2}{2m} - \epsilon_F \tag{5.7}$$

---

[11]Which we are told all the time in QM we are allowed to do; note that really this is a lie, but let us not worry about it.

if we do this and plug this into the energy, we get something nicer, as both particles and holes will have positive energy relative to the vacuum state (we return to this statement shortly). The net effect on the Hamiltonian is the substitution:

$$H \rightarrow H - \epsilon_F \mathcal{N}. \tag{5.8}$$

Now, if we want to calculate the number operator, we have:

$$\mathcal{N} = \int_{|\mathbf{k}|>k_F} d^3k a^{\dagger\sigma}(\mathbf{k}) a_\sigma(\mathbf{k}) - \int_{|\mathbf{k}|<k_F} d^3k b^\dagger_\sigma(\mathbf{k}) b^\sigma(\mathbf{k}) + \rho V \tag{5.9}$$

The negative sign on the b integral comes from the fact that we interchange the order to get the correct order of operators, but then we get a negative sign from the anticommutation, and we add ρV (with ρ the density and V the volume... an infinite constant. It is counting all of the particles in the Fermi sea/inside the Fermi sphere). If we now write down the Hamiltonian, we have:

$$H - \epsilon_F \mathcal{N} = \int_{|\mathbf{k}|>k_F} d^3k \frac{\hbar^2(\mathbf{k}^2 - k_F^2)}{2m} a^{\dagger\sigma}(\mathbf{k}) a_\sigma(\mathbf{k}) - \int_{|\mathbf{k}|<k_F} \frac{\hbar^2(k_F^2 - \mathbf{k}^2)}{2m} b^\dagger_\sigma(\mathbf{k}) b^\sigma(\mathbf{k}) + \phi V \tag{5.10}$$

where the change in the sign of the energy is from the anticommutation of the bs, and we also add the energy (an infinite constant representing the energy of all particles within the Fermi sphere) to compensate for this anticommutation. Where:

$$\Phi = \langle H - \epsilon_F N \rangle = \phi V \tag{5.11}$$

is the Grand canonical free energy. Note if we open our system, the fermions were go out of the system until none are left (they tend to want to escape to minimize the free energy); in order for this to not happen, there is a chemical potential μ which can be seen as a sort of "gate voltage" of the fermions. It is an adjustable parameter we can tweak to coax the fermions back in. So then, determining the density has to do with extremizing the free energy.

Having do this, we notice that the energy of our excitations are all positive! So our shifting of the Hamiltonian was a good choice.

5.3 Equations of State and Physical Parameters

How do we figure out the parameters in the above expression (e.g. the density)? Let's return to the particle number in our old formalism:

$$\mathcal{N} = \int d^3k \alpha^{\dagger\sigma}(\mathbf{k}) \alpha_\sigma(\mathbf{k}) \tag{5.12}$$

From this, let's calculate the expectation value for the vacuum state[12]:

$$\rho V = \langle \mathcal{O} | \mathcal{N} | \mathcal{O} \rangle = \int_{|\mathbf{k}|<k_F} (2J+1) \delta^3(\mathbf{0}) \tag{5.13}$$

where we have noted that the integral vanishes when $|\mathbf{k}| > k_F$ due to how the vacuum is defined, and we have replaced the αs with a terrible anticommutator. Here $\delta^3(\mathbf{k})$ is defined as:

$$\delta^3(\mathbf{k}) = \int \frac{d^3x}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{x}} \tag{5.14}$$

so then:

$$\delta^3(\mathbf{0}) = \int \frac{d^3x}{(2\pi)^3} = \frac{V}{(2\pi)^3} \tag{5.15}$$

---

[12]Note we will be very mathematically relaxed in our manipulations here... e.g. interchanging orders of matrix elements. It may be possible to justify this with working in a finite system and then taking limits, but we get the same answer so let us be cavalier about it.

so the above expectation value becomes:

$$\rho V = \int_{|\mathbf{k}|<k_F} (2J+1)\frac{V}{(2\pi)^3} \tag{5.16}$$

and if we now cancel out the volume, we remove the infinite quantities from both sides:

$$\rho = \int_{|\mathbf{k}|<k_F} \frac{2J+1}{(2\pi)^3} = \frac{(2J+1)}{(2\pi)^3}\frac{4\pi}{3}k_F^3 = \frac{(2J+1)}{6\pi^2}k_F^3 \tag{5.17}$$

where $\frac{4\pi}{3}k_F^3$ is the volume contained in the Fermi-sphere of radius k_F. Another calculation we can do is to calculate the internal energy; this is just the expectation value of our original Hamiltonian, which works out in a very similar way:

$$U = \langle \mathcal{O}|H|\mathcal{O}\rangle = V\frac{(2J+1)}{(2\pi)^3}\int_{|\mathbf{k}|<k_F} d^3k \frac{\hbar^2\mathbf{k}^2}{2m}. \tag{5.18}$$

We go to polar coordinates to calculate the above:

$$U = V\frac{(2J+1)}{(2\pi)^3}\frac{\hbar^2}{2m}\int_0^{k_F} k^2 dk \int_0^\pi \sin(\theta)d\theta \int_0^{2\pi} d\varphi k^2 = V\frac{(2J+1)}{(2\pi)^3}\frac{\hbar^2}{2m}\left(\frac{k_F^5}{5}\right)(4\pi) \tag{5.19}$$

So we conclude:

$$U = V\frac{(2J+1)}{(2\pi)^2}\frac{\hbar^2}{m}\frac{k_F^5}{5} = uV \tag{5.20}$$

where u is the energy density. So, we can now use these two results to solve for k_F and then eliminate it from our expressions. This gives us equations of state, which contain valuable (nontrivial!) physical information. We find from Eq. (5.17):

$$k_F = \left[\frac{6\pi^2\rho}{(2J+1)}\right]^{1/3} \tag{5.21}$$

and so plugging into Eq. (5.20) we have:

$$u = \frac{\hbar^2}{2m}\frac{(2J+1)}{10\pi^2}\left[\frac{6\pi^2\rho}{(2J+1)}\right]^{5/3} \tag{5.22}$$

so $u \propto \rho^{5/3}$. We can then find various other physical quantities by using thermodynamic relations, e.g. the pressure as a function of energy density (note we have assumed that everything is at constant zero temperature here):

$$P = -\left.\frac{\partial U}{\partial V}\right|_N = \frac{2}{3}u \tag{5.23}$$

We can also find the chemical potential:

$$\mu = \left.\frac{\partial U}{\partial N}\right|_V = \epsilon_F \tag{5.24}$$

where in the above calculation we would write $U = uV$, and then write $\rho = N/V$. So we find at zero temperature that indeed the chemical potential is the Fermi energy (of course if one turns on interactions, or the temperature, then this will no longer hold!) What's more, if we look at Φ, we have:

$$\Phi = \langle H - \mu\mathcal{N}\rangle = \phi V \implies \phi = -P \tag{5.25}$$

5.4 A Simple Example of a Quantum Phase Transition

Note that there is a more detailed discussion of the above in the notes, and we have covered the main points here. But let us discuss one thing that is not in the notes. We recall the interplay between the fermions wanting to get away and the chemical potential to coax them back in. We now consider a simple model of a *quantum phase transition*[13]. We assume that we can tune the chemical potential; as we tune, the density varies. If $\mu = 0$, the density is zero. If we make it negative, the density stays at zero. So, as we increase μ, the density becomes nonzero as we cross $\mu = 0$; signalling a phase transition in the system.
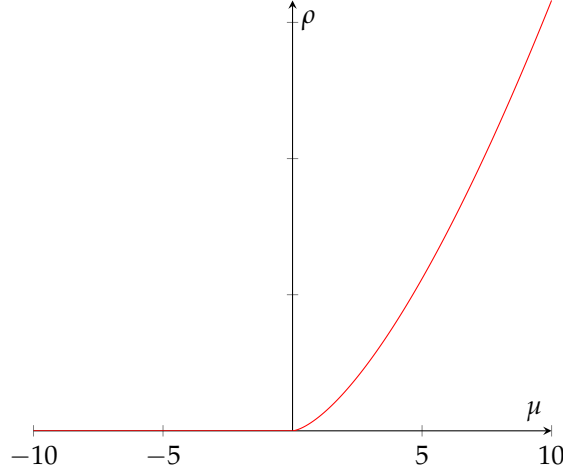
Figure 5.1: Plot of ρ vs. μ (arbitrary units). ρ is zero for $\mu \leq 0$ and $\rho \propto \mu^{3/2}$ for $\mu > 0$, signalling a phase transition in the system.

The fact that $\rho \propto \mu^{3/2}$ is obtainable by realizing that $k_F \propto \rho^{1/3}$ (Eq. (5.21)) and $\mu = \epsilon_F \propto k_F^2$ (Eqs. (5.24) and the definition of Fermi energy). Since $\varphi \propto P \propto u$ (Eqs. (5.25) and (5.23)), and $u \propto \rho^{5/3}$ (Eq. (5.17)), we find that the free energy φ is related to the chemical potential by $\varphi \propto \mu^{5/2}$.

Since we have to take three derivatives of φ w.r.t. μ before we get a discontinuity (at $\mu = 0$), by the Ehrenfest classification of phase transitions, this is a third order phase transition. This discussion can also be taken as something that highlights the difference between Fermi gas and Fermi liquid behaviour; interactions change the properties of the system significantly.

---

[13]One of the main models discussed in Sachdev's book by this title!

6 Degenerate Bose Gas

6.1 Setting up the Boson Case

Last time we discussed the Fermi gas, though we did not introduce any interactions. This is a starting point of doing condensed matter physics, nuclear physics (where the nucleus could be thought of as a small Fermi gas; though of course the interactions are strong in this case), neutron stars etc. When we add interactions, we get a Fermi liquid, where calculations become intractable; though we can apply computational methods in some cases. In a way, the strong interaction has been solved via a computer calculation; this is impressive as the models have been shown to work! Today, we instead discuss the degenerate Bose gas. We cover it as the second topic as it is more complex; due to Bose-Einstein condensation. We're stuck at zero temperature, and a gas of quantum-mechanical bosons at $T = 0$ (indeed, below some critical temperature T_c) forms a BE. If we're looking at particles without charge, we're looking at some kind of *superfluid*. The noninteracting case is so ugly that we don't study it; we add a small interaction to make things more sensible, so the use of "fluid" is really correct, here.

We study a zero-temperature state of a box of bosons. If we ignore all interactions, the energy takes the form:

$$\epsilon = \frac{\hbar^2 \mathbf{k}^2}{2m}. \tag{6.1}$$

Here, more than one particle can occupy the lowest energy state, so we might take our quantized field with the creation operator, and do something that looks nonsensical:

$$(\alpha^\dagger(0))^N |0\rangle \tag{6.2}$$

Of course this looks crazy, as these particles with $\mathbf{k} = \mathbf{0}$ have infinite wavelength... so perhaps we should do something else, but what else is there? We could say that there is confinement within a box (of box L with hard boundary conditions), but then the wavefunction would follow boundary conditions (in each direction).
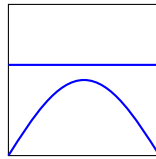
Figure 6.1: Two possible views as the problem. We can constrain the bosons to a box of side length L with hard boundaries, constraining the ground state wavefunction to look like $\sqrt{\frac{2}{L}} \sin(\frac{2\pi x}{L})$. Alteratively, we can assume our system is a finite patch out of infinite space, in which case the ground state wavefunctions are plane waves with infinite wavelength, i.e. constant in space.

In fact, the wavefunction and the energy would be quite sensitive to the boundary conditions. This is not something that we really want[14]. We could do the opposite and remove the walls, and assume that our system is a finite patch within infinite space. Then a piece of the system should look fairly generic. But what happens then? The wavefunction would be constant over all space if $\mathbf{k} = \mathbf{0}$, so the probability of finding it anywhere in space is equal. So, our state is not something with a constant particle number. So at the outset, we can take this point of view that the particle number is not fixed; an approach that particle physicists and field theorists love. One could instead take a closed system and study it through this method, but there is not a consensus as to whether the two approaches are actually equivalent. But we will take the field theorists approach, here.

---

[14]as theorists in a QFT class, anyway; maybe not if you work at SBQMI down the street.

Motivated by this argument, we consider the state:

$$|\mathcal{O}\rangle = \sum_{N=0}^{\infty} c_N \left(\alpha^\dagger(\mathbf{0})\right)^N |0\rangle. \tag{6.3}$$

This is still not quite sensible; perhaps we can take something like:

$$\alpha_f^\dagger = \int d^3x f(\mathbf{x}) \alpha^{dag}(\mathbf{k}) \tag{6.4}$$

for f square integrable if we wanted to be a tad more rigorous, but let us not worry ourselves with this very much. How do we characterize such states? One thing we can notice is:

$$\langle \mathcal{O}|\psi(\mathbf{x},t)|\mathcal{O}\rangle \neq 0. \tag{6.5}$$

The field operator annihilates a boson, but it will still have a nonzero overlap with the original vacuum state, so the overall expectation value will be nonzero (see HW2). Note that this is okay for bosons, but we would never see this for fermions. This is hugely degenerate if the bosons are non-interacting.

6.2 Coherent States

Let us talk for a little about coherent states, which have the above expectation value property. At $t = 0$, we consider the field-operator commutation relations:

$$[\psi(\mathbf{x}), \psi^\dagger(\mathbf{y})] = \delta^3(\mathbf{x} - \mathbf{y}) \tag{6.6}$$

and:

$$[\psi^\dagger(\mathbf{x}), \psi^\dagger(\mathbf{y})] = [\psi(\mathbf{x}), \psi(\mathbf{y})] = 0. \tag{6.7}$$

If we act the annhilation operator on the empty vacuum, we have:

$$\psi(\mathbf{x})|0\rangle = 0. \tag{6.8}$$

We do not concern ourselves with the spin. Now, we consider the state:

$$|\eta\rangle = e^{\int d^3x \left(\eta^*(\mathbf{x})\psi(\mathbf{x}) - \psi^\dagger(\mathbf{x})\eta(\mathbf{x})\right)}|0\rangle \tag{6.9}$$

where the operator in the exponential is anti-hermitian and so the overall operator is unitary. If we write the normal ordered version of the above, we obtain:

$$|\eta\rangle = e^{-\int d^3x \eta^*(\mathbf{x})\eta(\mathbf{x})} e^{-\int d^3x \eta(\mathbf{x})\psi^\dagger(\mathbf{x})}|0\rangle \tag{6.10}$$

where we have got rid of the ψ term by considering that this annhilates the vacuum state. We can write the action of the annihilation operator on the coherent state as:

$$\psi(\mathbf{x})|\eta\rangle = \eta(\mathbf{x})|\eta\rangle. \tag{6.11}$$

So that's a cool property! It's also a unitary transform of the vacuum state, so it is normalized:

$$\langle \eta|\eta\rangle = 1. \tag{6.12}$$

So this is an example of the state where:

$$\langle \psi|\psi(\mathbf{x})|\eta\rangle = \eta(\mathbf{x}). \tag{6.13}$$

This is an example of a "good" coherent state (as opposed to a bad one) and very often in physics we used bad ones. What could go wrong? For example, the integral $\int d^3x \eta^*(\mathbf{x})\eta(\mathbf{x})$ could diverge if $\eta(\mathbf{x})$ is a constant. This often happens, actually (and we will proceed to work with this right now). The bad coherent state is ubiquotous; any interaction state of charges particles produces a "coherent" state of soft photons, but this is a bad coherent state (it has no overlap with any state which has a finite number of photons). Every QED interaction produces an infinite number of photons, which fly away undetected, but seem to be always there. So, we drop "true/good" coherent states for now, and work with bad ones; which will still give us sensible results.

6.3 Landau's Argument for Superfluidity

Note that in the textbook that there is a section which reviews Landau's famous argument about the quasiparticle spectrum and critical velocity of a superfluid. It uses Galilean symmetry, and so we don't cover it here (and the treatment of the book is so refined, that it is probably better to read it there). There is however the idea of a superfluid flowing through a pipe, without resistance, and so even if we introduce some interactions (between the particles, and with the pipe), it will flow through the pipe without any resistance. This is to some extent seen in the lab, and there is seen that if the superfluid goes too fast then the superfluid starts to feel resistance. Landau's argument goes as follows. For the energy to dissipate, there should be some excitations in the fluid. So, there should be some viscosity that creates a travelling wave/ripple (sometimes called quasiparticles) in the fluid. Let us say this is wavelike (everything is quantum mechanical here), with a wavenumber $\mathbf{k}$ and frequency $\omega(\mathbf{k})$.
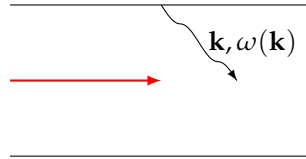
Figure 6.2: In Landau's argument for superfluids, we consider a superfluid travelling in a pipe, and wavelike excitations/quasiparticles with wavenumber $\mathbf{k}$ and frequency $\omega(\mathbf{k})$.

He then argues that this should be energetically favourable if the velocity is larger than a critical velocity v_c:

$$v_c = \min_{\mathbf{k}} \frac{\omega(\mathbf{k})}{|\mathbf{k}|}. \tag{6.14}$$

The argument is beautiful in how it invokes Galilean relativity/symmetry (do read about it in your own time!) Note the stark difference from the free particle case; as then we would have $\omega(\mathbf{k}) \propto \mathbf{k}^2$ so $v_c = 0$, i.e. free particles are not superfluidic. What does work is if we have a sound wave, as then $\omega(\mathbf{k}) = v_s|\mathbf{k}|$; here the critical velocity is the speed of sound.

So, the important goal for us; can we find this using our model? What do the elementary excitations look like?

A small aside; this argument doesn't really agree with experiments. v_c as measured is smaller than what we would expect from Landau's argument. The explanation is that the dispersion relation $\omega(\mathbf{k})$ as measured in experiment takes the form of the roton curve as seen in the below figure. Then v_c is not the slope of the purple curve (as would be predicted theoretically) but instead is the slope of the red curve (less than the theoretically predicted value).

6.4 Introducing a Small Interaction to our QFT

We introduce a weak repulsive interaction between the bosons that supresses the Bose-Einstein condensation. They will want to be far apart and run away from our system; so we add a chemical potential to balance this and draw them back in (so we end up with a finite density at the end). Let us write down a Hamiltonian[15] for this:

$$H = \int d^3x \left(\frac{\hbar^2}{2m} \boldsymbol{\nabla}\psi^\dagger(\mathbf{x}) \cdot \boldsymbol{\nabla}\psi(\mathbf{x}) - \mu\psi^\dagger(\mathbf{x})\psi(\mathbf{x}) + \frac{\lambda}{2}\psi^\dagger(\mathbf{x})\psi^\dagger(\mathbf{x})\psi_\mathbf{x}\psi(\mathbf{x}) \right) \tag{6.15}$$

where the last term is a ball-bearing potential, with λ "small" (and positive; else the spectrum will not be bounded from below, as the energy from all the particles sitting on top of each other will be negative

---

[15]A remark: model building is often much easier than the actual toying with the (e.g. field) equations...

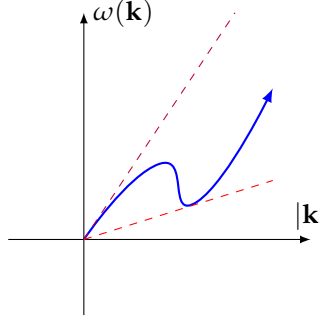Figure 6.3: Plot of the experimental dispersion relation for superfluids (roton dispersion relation). From this we can see the reason for the experimental disagreement of the Landau argument for determining the critical velocity v_c of superfluids, as the local minimum in the roton curve shifts the critical velocity to be the slope of the red curve, rather than the Landau theoretical prediction of the purple curve.

infinity). We also enforce the commutation relations:

$$[\psi(\mathbf{x}, t), \psi^\dagger(\mathbf{y}, t)] = \delta^3(\mathbf{x} - \mathbf{y}), \quad [\psi^\dagger(\mathbf{x}, t), \psi^\dagger(\mathbf{y}, t)] = [\psi(\mathbf{x}, t), \psi(\mathbf{y}, t)] = 0 \tag{6.16}$$

and also introduce the vacuum state $|\mathcal{O}\rangle$, which is the lowest-eigenvalue eigenstate of the (bounded-from-below) Hamiltonian. It is possible that the expectation value is nonvanishing, so let us anticipate this and introduce a function $\eta(\mathbf{x})$:

$$\langle \mathcal{O} | \psi(\mathbf{x}, t) | \mathcal{O} \rangle = \eta(\mathbf{x}) \neq 0. \tag{6.17}$$

How do we treat this possibility in a systematic way? One way would be to write:

$$\psi(\mathbf{x}, t) = \eta(\mathbf{x}, t) + \tilde{\psi}(\mathbf{x}, t) \tag{6.18}$$

where η is the classical part and $\tilde{\psi}$ is the quantum/non-classical part, i.e. which follows:

$$\langle \mathcal{O} | \tilde{\psi}(\mathbf{x}, t) | \mathcal{O} \rangle = 0. \tag{6.19}$$

So, there will be a classical part of the Hamiltonian where we forget about the $\tilde{\psi}$s, and in this term we would like ηs which minimize the Hamiltonian. Then we can solve a classical field equation, check the extremum of the functional etc. We can argue that the η that minimizes the Hamiltonian goes as:

$$\eta \sim \frac{1}{\sqrt{\lambda}}. \tag{6.20}$$

At very weak coupling, this classical piece is emphasized. We then might ask what the size of the quantum part is. Since ψ needs to obey equal-time commutation relations, and η drops out of these (as it is classical and drops out of the commutation relations), then since $\tilde{\psi}$ follows the commutation relations (where λ shows up nowhere), this tells us that $\tilde{\psi}$ is of order unity:

$$\tilde{\psi} \sim \lambda^0 = 1. \tag{6.21}$$

If we make λ small, we can then ignore anything but the classical η term. We can consider an asymptotic expansion of η (asymptotic as the first term is singular) and $\tilde{\psi}$, each of which have corrections which can be solved order by order. This is an arbitrarily good approach so long as λ is arbitrary small and nonzero. Let us take it (next class!)

Another thing we might be interested in is the "smoothest" possible states, i.e. states that have some symmetry. A smooth surface has translation symmetry. We expect that the low-energy states of our theory

24

here may be smooth. Suppose η was a constant; then the most important part of our field is smooth. This is like going back to our condensate at the start of our lecture, where our wavefunctions were constant throughout all space. We then simply search for a constant which minimizes the Hamiltonian. The kinetic energy terms are zero if η is a constant, and we can minimize the other terms to find:

$$\eta = \sqrt{\frac{\mu}{\lambda}} \sim \frac{1}{\sqrt{\lambda}}. \tag{6.22}$$

From our theory we can calculate the density:

$$\rho \cong \frac{\mu}{\lambda} \tag{6.23}$$

and the (grand canonical free) energy:

$$\Phi = -\frac{\mu^2}{2\lambda} + \ldots = -\frac{\lambda}{2}\rho^2 + \ldots \tag{6.24}$$

which is also equal to the negative of the pressure:

$$P = -\Phi = \frac{\lambda}{2}\rho^2. \tag{6.25}$$

So already this theory gives us a lot to work with, but we have yet to show that it is in fact a superfluid; this is something we will calculate and verify next class, by figuring out what $\omega(\mathbf{k})$ is for our theory.

7 Degenerate Bose Gas Continued

7.1 A Review of the Boson Gas Hamiltonian

Recall the Hamiltonian we were working with in studying the Bose gas/liquid:

$$H = \int d^3x \left(\frac{\hbar^2}{2m} \boldsymbol{\nabla}\psi^\dagger \cdot \boldsymbol{\nabla}\psi - \mu\psi^\dagger\psi + \frac{\lambda}{2}\psi^\dagger\psi^\dagger\psi\psi \right) \tag{7.1}$$

where:

$$\psi = \eta + \tilde{\psi} \tag{7.2}$$

with η is the classical part and $\tilde{\psi}$ the quantum part. We can interpret this as:

$$\langle \mathcal{O}|\tilde{\psi}|\mathcal{O}\rangle = 0. \tag{7.3}$$

We noted that the classical part η was very important in the weak-coupling limit, as $\eta \sim \frac{1}{\sqrt{\lambda}}$. Meanwhile, $\tilde{\psi}$ obeys commutation relations for which no λ shows up, so $\tilde{\psi} \sim \lambda^0$. However, we expect that there are an infinite series of correction, so really:

$$\psi = \eta + \tilde{\psi} + \delta\eta + \lambda\delta\tilde{\psi} \tag{7.4}$$

Where $\delta\eta \sim \sqrt{\lambda}$. For small λ (though this is not entirely a trivial statement; λ has dimensions, so what it is small compared to?), it is meaningful to analyze the leading terms (i.e. a classical Hamiltonian). We can plug in η to where the ψs are in the Hamiltonian, and we get something that looks like a potential:

$$V(\eta) = \mathcal{V}(-\mu|\eta|^2 + \frac{\lambda}{2}|\eta|^4) \tag{7.5}$$

apologies for the confusing notation; V on the left is the potential, $\mathcal{V}$ on the right is a volume. Minimizing V (the potential) with respect to η, we find:

$$\eta = \sqrt{\frac{\mu}{\lambda}}. \tag{7.6}$$

There is something funny with this identification; $V(\eta)$ depends only on the norm of η, but in the expression for η we have chosen it to be real. From a minimization perspective:

$$\eta = \sqrt{\frac{\mu}{\lambda}}e^{i\theta} \tag{7.7}$$

are valid minima for all $\theta \in \mathbb{R}$. So the minimum is not unique. However, we can proceed by just choosing an angle of our choice, and it will not matter. Why is this the case?

7.2 A Brief Foray Into Spontaneous Symmetry Breaking

Looking at the Hamiltonian, we notice that there is a symmetry; namely, the Hamiltonian is unchanged by the introduction of some phase $\psi \mapsto \psi e^{i\theta}$. So, this tells us that what we choose for the phase of η should not matter.

The idea is that the potential is completely symmetric under rotation, but the solution is not. This is something known as spontaneous symmetry breaking. Classical analogy: which way does chalk fall when one lets go of it from the top position. A priori, there is a symmetry; there is no preference to how it falls. But when we actually do the experiment, it falls somewhere as we perturb it somehow when releasing (breaking the symmetry) or an air molecule bounces off of it and causes it to fall. Quantum mechanically, the chalk is described by the rigid rotator Hamiltonian:

$$H = \frac{I}{2}\dot{\theta}^2 \tag{7.8}$$
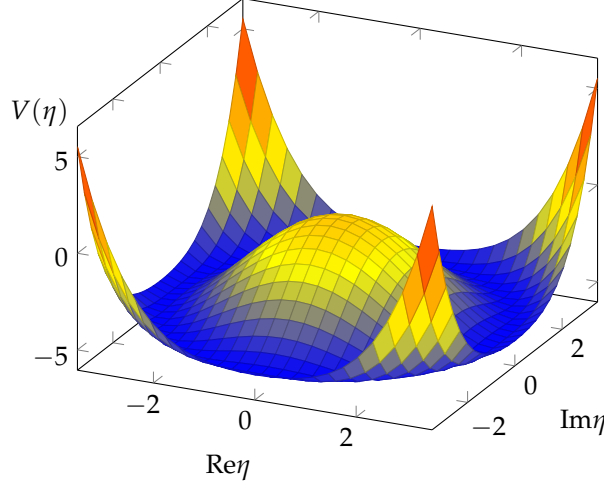
Figure 7.1: 3D plot of the potential V as a function of η. We take $\mathcal{V} = 1, \mu = 1, \lambda = 1/10$. The blue ring corresponds to $\eta = \sqrt{\frac{\mu}{\lambda}} e^{i\theta}$ which minimizes the potential, which is radially symmetric.

If I look at the eigenfunctions, we have:

$$\psi(\theta) = e^{in\theta} \tag{7.9}$$

The ground state is a superposition of all of these angles, which averages out to zero. But this is clearly not the right state; it is in some orientation and it is frozen there. The idea is that the chalk pointing at some angle is actually an excited state; we have something like:

$$\psi(\theta) = \sum_n e^{in(\theta - \theta_c)} \tag{7.10}$$

and it relaxes towards the ground state given sufficient time; however this time depends on the moment of inertia, which here it is very large. If the chalk was small enough, then we would see it undergo this relaxation; since it is macroscopic, the relaxation time is very large.

Why this chalk discussion? This is to say that why the quantum mechanical description of the chalk may be correct, it is really not particularly reasonable; it makes more sense to treat this classically. So, returning back to the idea of spontaneous symmetry breaking. This is a ubiqitous phenomenon (in condensed matter, particle physics, etc.). For example, consider ferromagnetism. The true ground state of a ferromagnetic is some superposition, but if we look at it, it will not decay to the ground state during the lifetime of the universe. So, we can just think about this classically.

Having now chosen a solution, we can't see the symmetry of the whole potential anymore; the fluctuations (as obtained by calculating the higher order terms) cannot "see" the whole energy landscape (one can think that the "moment of inertia" is so large that it cannot explore the entire ring of minima).

A takeaway from all of this; spontaneous symmetry-breaking is a phenomena we often see when classical effects dominate over quantum ones.

7.3   Solving the System

So, we've solved for the leading term of the expansion of ψ! Since it's dominant, we could just calculate the internal energy, or the grand canonical free energy from what we have so far. We can find the GC free energy to be:

$$\phi = -\frac{\lambda}{2}\rho^2 = -P. \tag{7.11}$$

27

We can find corrections to this by calculating the contributions from the higher order terms (these corrections should be small, if our asymptotic expansion is actually reasonable). In particular, here this expansion works at low density.

Next, we can plug $\psi = \eta + \tilde{\psi}$ into the field equation (see HW2) or the Hamiltonian (which we do now), and see what we get out. We find:

$$H' = -\frac{\lambda}{2}\rho^2 V + \int d^3x \left(\frac{\hbar^2}{2m}\boldsymbol{\nabla}\tilde{\psi}^\dagger \cdot \boldsymbol{\nabla}\tilde{\psi} + \mu\tilde{\psi}^\dagger\tilde{\psi} + \frac{\mu}{2}\tilde{\psi}\tilde{\psi} + \frac{\mu}{2}\tilde{\psi}^\dagger\tilde{\psi}^\dagger \right) + \mathcal{O}(\lambda) \tag{7.12}$$

Now the Hamiltonian does not look normal in the $\tilde{\psi}$ and $\tilde{\psi}^\dagger$s anymore; this is due to the broken symmetry. We can also plug in our expansion into the commutation relations:

$$[\tilde{\psi}(\mathbf{x},t), \tilde{\psi}^\dagger(\mathbf{y},t)] = \delta^3(\mathbf{x}-\mathbf{y}) \tag{7.13}$$

Let us comment on the $\mathcal{O}(\lambda)$ corrections; we first have that:

$$-\frac{\lambda}{2}\rho^2 V = -\frac{\mu^2}{2\lambda}V \tag{7.14}$$

So this is indeed the leading term for small λ. The integral term is of order $\lambda^0 = 1$. Then the corrections are of order λ, so we can control these by taking small λ.

For large derivatives, the kinetic energy terms dominate, but for smaller derivatives the other terms in the integral become important.

By plugging things back in, we get a linear field equation for $\tilde{\psi}$ (The Bogolibov-de-Gennes equation); see HW2. But let us attack this Hamiltonian directly:

$$\psi(\mathbf{x},0) = \int \frac{d^3k}{(2\pi)^{3/2}} e^{i\mathbf{k}\cdot\mathbf{x}}\alpha(\mathbf{k}) \tag{7.15}$$

this is really just a change of basis in function space; the whole integral is like a unitary matrix that rotates the basis. Let us go to the αs; why? Because the Hamiltonian is translation invariant, we would suspect that the momentum would be an important quantity, and $\mathbf{k}$ is closely related to momentum, so it is good to analyze. The Hamiltonian reads:

$$H = -\frac{\mu^2}{2\lambda}V + \int d^3k \left(\frac{\hbar^2 k^2}{2m}\alpha^\dagger(\mathbf{k})\alpha(\mathbf{k}) + \mu\alpha^\dagger(\mathbf{k})\alpha(\mathbf{k}) + \frac{\mu}{2}\alpha(-\mathbf{k})\alpha(\mathbf{k}) + \frac{\mu}{2}\alpha^\dagger(\mathbf{k})\alpha^\dagger(-\mathbf{k}) \right) \tag{7.16}$$

From going to k-space, we have a separate Hamiltonian for each value of $\mathbf{k}$, and the first term ($\alpha^\dagger(\mathbf{k})\alpha(\mathbf{k})$) is diagonalized. However, it is still not clear what the eigenvalues of the elementary excitations are. We need to diagonalize this Hamiltonian; let us try this. First we note that the Hamiltonian is of quadratic form, so we should be able to diagonalize it. We let:

$$\alpha(\mathbf{k}) = \cosh\theta(\mathbf{k})a(\mathbf{k}) - \sinh\theta(\mathbf{k})a^\dagger(-\mathbf{k})$$
$$\alpha^\dagger(\mathbf{k}) = \cosh\theta(\mathbf{k})a^\dagger(\mathbf{k}) - \sinh\theta(\mathbf{k})a(-\mathbf{k}) \tag{7.17}$$

Wait, where do the hyperbolic functions come from? We know that the αs follow the commutation relations:

$$[\alpha(\mathbf{k}), \alpha^\dagger(\mathbf{l})] = \delta^3(\mathbf{k}-\mathbf{l}) \tag{7.18}$$

so we want the as to do the same:

$$[a(\mathbf{k}), a^\dagger(\mathbf{l})] = \delta^3(\mathbf{k}-\mathbf{l}) \tag{7.19}$$

I don't have a lot of freedom in the coefficients if I want the commutation relations to hold; the hyperbolic functions come out of this.

Eqs. (7.17) is known as a Bogoliubov transformation. We can plug this into the Hamiltonian and get an even more complicated expression with $\sin\theta(\mathbf{k})$s and as, and we can adjust $\theta(\mathbf{k})$ until we get something proportional to $a^\dagger(\mathbf{k})a(\mathbf{k})$. When we do so, we end up with:

$$H = -\frac{\mu^2}{2\lambda}V + \int d^3k E(\mathbf{k})a^\dagger(\mathbf{k})a(\mathbf{k}) + E_0 \tag{7.20}$$

where the constant at the end comes from the corrections from interchanging the order of operators (as a consistency check: this constant should be small compared to the leading order constant). It is a kind of zero-point energy. In order for the above to hold, we find the constraint that:

$$\tanh 2\theta(\mathbf{k}) = \frac{\mu}{\frac{\hbar^2\mathbf{k}^2}{2m} + \mu} \tag{7.21}$$

and $E(\mathbf{k})$ follows:

$$E(\mathbf{k}) = \sqrt{\left(\frac{\hbar^2\mathbf{k}^2}{2m} + \mu\right)^2 - \mu^2} \tag{7.22}$$

so we obtain a dispersion relation for $\mathbf{k}$! The cool thing is we can now use this dispersion relation combined with the Landau result to get a testable prediction of the critical velocity for superfluids. For small $\mathbf{k}$, the above reduces to:

$$E(\mathbf{k}) \approx \sqrt{\frac{\hbar^2}{2m}\mu}|\mathbf{k}| \tag{7.23}$$

So the Landau criterion would tell us that:

$$v_c = \sqrt{\frac{\hbar^2}{2m}\mu}. \tag{7.24}$$

This turns out to be wrong most of the time, as superfluids are in general not weakly interacting. But, when an experimental group created a weakly-interacting Bose-Einstein condensate, they were able to indeed verify that the above expression holds!

This concludes our discussion; HW2 attacks the same problem from the perspective of the equations of motion.

7.4 A Small Teaser for Next Lecture

We've developed our QFT, rewrote many-particle QM as a QFT, and studied some simple (but useful!) examples (such as the Fermi gas[16] and superfluids). Now, we go into developing the ideas of QFT a little further. One might wonder why we need something else; this is because this "something else" gives us great power and a nicer way to study field theories when we have to do perturbation theory. It may be an unexpected step, but we will temporarily regress from doing QFT to doing classical field theory. We will consider the *action principle*. We will consider some classical field $\psi_\sigma(\mathbf{x}, t)$ (bosons only; anti-commuting fermions would not be pure functions, but would instead be operators belonging to an algebra...). We will then ask if we can derive our field equation from a variational principle and an action. Of course the answer is yes (else we wouldn't bring it up), and we can even write down the Lagrangian density and the action which will yield our classical field equation. The action is:

$$S = \int dt d^3x \mathcal{L} \tag{7.25}$$

---

[16]Quite a useful "first-order" approximation for metals.

where the Lagrangian density is:

$$\mathcal{L} = i\hbar \psi^{\dagger\sigma} \frac{\partial}{\partial t} \psi_\sigma - \frac{\hbar^2}{2m} \boldsymbol{\nabla} \psi^{\dagger\sigma} \cdot \boldsymbol{\nabla} \psi_\sigma + \mu \psi^{\dagger\sigma} \psi_\sigma - \frac{\lambda}{2} \left(\psi^{\dagger\sigma} \psi_\sigma \right)^2. \tag{7.26}$$

However, we have lost something: our quantum field equation had a particular ordering of the ψ and $\psi^\dagger$s in the interaction part. We have to recover this somehow. This could be a small or large problem; for us it is small because we know how it has to look like ahead of time. But, something like the ordering problem in treating Einstein's gravity is not resolved. In general, this is something one needs to solve case-by-case.

What then, do we gain? The Lagrangian encodes the commutator as well as the field equation. One can encode the first term as:

$$\left(i\hbar \psi^{\dagger\sigma} \right) \left(\frac{\partial}{\partial t} \psi_\sigma \right) = P\dot{Q} \tag{7.27}$$

and note the connection with the Hamiltonian:

$$L = P\dot{Q} - H(P, Q). \tag{7.28}$$

One can consider the Poisson bracket:

$$\{\psi_\sigma(\mathbf{x}, t), i\hbar \psi^{\dagger\rho}(\mathbf{y}, t)\} = \delta^\rho_\sigma \delta^3(\mathbf{x} - \mathbf{y}) \tag{7.29}$$

and when quantizing one adds an $i\hbar$ and replaces the Poisson bracket with a commutator. So there is that benefit of less writing. We also gain the analysis of symmetries and conserved quantities that we know from classical field theory, which we can lift to the QFT (for example, conservation of energy, momentum).

This is not just a construct; we will learn how to use functional integrals, and the (classical) action plays a very large role there (we integrate over it). It will stay with us for a long time.

8 The Action Principle

We revert a little bit back into classical field theory. Technically, the field equation and the commutation relations we have constructed define our QFT, and give us all the information we need. So changing the formalism needs some justification; that justification is that classical field theory has a lot of structure, we can lift back into QFT to learn more about QFT. This logic is fairly obscure in most textbooks, but we will explore it a bit deeper. We will find the classical to quantum transition to be relatively straightforwards. For generic quantum theories, this is not the case, but quantum field theories tend to be simple as far as mechanical things go. This is not just because we like simple things, but because the internal consistency of quantum field theories doesn't let them be extremely complicated.

8.1 The Action & Lagrangians

The action is defined as the following integral over spacetime:

$$S = \int d^3x dt \mathcal{L} = S[\psi, \psi^\dagger] \tag{8.1}$$

where $\mathcal{L}$ is the Lagrangian density. It is a functional; it maps functions $\psi, \psi^\dagger$ onto a number S. We consider the Lagrangian density:

$$\mathcal{L} = \frac{i\hbar}{2}\psi^{\dagger\sigma}\frac{\partial\psi_\sigma}{\partial t} - \frac{i\hbar}{2}\frac{\partial\psi^{\dagger\sigma}}{\partial t}\psi_\sigma - \frac{\hbar^2}{2m}\boldsymbol{\nabla}\psi^{\dagger\sigma}\cdot\boldsymbol{\nabla}\psi_\sigma + \mu\psi^{\dagger\sigma}\psi_\sigma - \frac{\lambda}{2}(\psi^{\dagger\sigma}\psi_\sigma)^2. \tag{8.2}$$

this is the Lagrangian that would recover our non-relativistic quantum field theory (we just conjured this up, with some guidelines; namely that it reproduces the field equation that we want. The field equation is the primitive here, and the Lagrangian we reverse engineer[17]). From this Lagrangian we can not only recover the field equation but also the commutator.

Note that for free field theories, we will in general only consider Lagrangians that are quadratic in the fields; this allows for the equations of motion to be linear. For interacting field theories we may have higher order terms.

Now, if we consider the relation of the Lagrangian to the Hamiltonian:

$$\mathcal{L} = P\dot{Q} - H(P,Q) = i\hbar\psi^{\dagger\sigma}\frac{\partial\psi_\sigma}{\partial t} - \dots \tag{8.3}$$

We can identify Q with $\psi_\sigma(\mathbf{x}, t)$ and P with $i\hbar\psi^{\dagger\rho}(\mathbf{x}, t)$. The Poisson bracket then gives us:

$$\{\psi_\sigma(\mathbf{x}, t), i\hbar\psi^{\dagger\rho}(\mathbf{x}, t)\} = \delta^\rho_\sigma\delta^3(\mathbf{x} - \mathbf{y}) \tag{8.4}$$

So if we then go from classical to quantum by changing the Poisson brackets to commutators and introducing a factor if $i\hbar$, we have:

$$[\psi_\sigma(\mathbf{x}, t), i\hbar\psi^{\dagger\rho}(\mathbf{x}, t)] = i\hbar\delta^\rho_\sigma\delta^3(\mathbf{x} - \mathbf{y}) \implies [\psi_\sigma(\mathbf{x}, t), \psi^{\dagger\rho}(\mathbf{x}, t)] = \delta^\rho_\sigma\delta^3(\mathbf{x} - \mathbf{y}). \tag{8.5}$$

Next, how do we get the field equation? The action principle states that the action functional, considered as a mapping of classical fields obeying the appropriate boundary conditions to the real numbers, is stationary when it is evaluated on the field configurations which obey the classical equations of motion, that is, the classical field equation. So, the Lagrangian encodes the field equation, and we can suss it out through some calculus of variations/functional calculus.

---

[17]But note that once we start to familiarize ourselves with Lagrangians, we will find it easier to write down a Lagrangian and derive the field equations etc. from there. It is a slightly shorter presentation of all of the data encoded in the field equation and commutation relations.

8.2 Deriving the Euler-Lagrange Equations

We write down a formal mathematical criteria for what this means. $\psi, \psi^\dagger$ are a "stationary point" (of course one remembers these are actually functions) when:

$$S[\psi + \delta\psi, \psi^\dagger + \delta\psi^\dagger] = S[\psi, \psi^\dagger] + O((\delta\psi)^2, (\delta\psi^\dagger)^2, \delta\psi\delta\psi^\dagger) = S[\psi, \psi^\dagger] + \delta S \tag{8.6}$$

One might ask what a "nearby point" in function space actually is (i.e. what does $\delta\psi$ mean?); we avoid this discussion as our applications tend to be simple.

We can now define the variation of the action as follows (by considering the variation of the Lagrangian, treating it as a function of $\psi_\sigma, \psi^{\dagger\sigma}$ and its spatial/time derivatives):

$$\delta S = \int dt d^3x \, \delta\mathcal{L}$$

$$= \int dt d^3x \left[\delta\psi_\sigma(\mathbf{x},t) \frac{\partial\mathcal{L}}{\partial\psi_\sigma(\mathbf{x},t)} + \delta\dot\psi_\sigma(\mathbf{x},t) \frac{\partial\mathcal{L}}{\partial\dot\psi_\sigma(\mathbf{x},t)} + \delta\boldsymbol{\nabla}\psi_\sigma(\mathbf{x},t) \frac{\partial\mathcal{L}}{\partial\boldsymbol{\nabla}\psi_\sigma(\mathbf{x},t)} + (\psi_\sigma \iff \psi^{\dagger\sigma}) \right] \tag{8.7}$$

where $(\psi_\sigma \iff \psi^{\dagger\sigma})$ denotes we have all the same terms, replacing the ψs with their Hermitian conjugate. We can then consider that:

$$\delta(\boldsymbol{\nabla}\psi_\sigma(\mathbf{x},t)) = \boldsymbol{\nabla}(\delta\psi_\sigma(\mathbf{x},t)) \tag{8.8}$$

i.e. the variation of the derivative is the derivative of the variation, and the same with the time derivatives:

$$\delta(\frac{\partial}{\partial t}\psi_\sigma(\mathbf{x},t)) = \frac{\partial}{\partial t}(\delta\psi_\sigma(\mathbf{x},t)) \tag{8.9}$$

We can then rewrite the variation of the action as:

$$\delta S = \int dt d^3x \left[\delta\psi_\sigma(\mathbf{x},t) \left(\frac{\partial\mathcal{L}}{\partial\psi_\sigma} - \frac{\partial}{\partial t}\frac{\partial\mathcal{L}}{\partial\dot\psi_\sigma} - \boldsymbol{\nabla}\cdot\frac{\partial\mathcal{L}}{\partial\boldsymbol{\nabla}\psi_\sigma} \right) + \frac{\partial}{\partial t}\left(\delta\psi_\sigma \frac{\partial\mathcal{L}}{\partial\dot\psi_\sigma} \right) + \boldsymbol{\nabla}\cdot\left(\delta\psi_\sigma \frac{\partial\mathcal{L}}{\partial\boldsymbol{\nabla}\psi_\sigma} \right) + (\psi_\sigma \iff \psi^{\dagger\sigma}) \right] \tag{8.10}$$

We look at the second and third terms in the above integral. It is a four-divergence, so Gauss's theorem allows us to rewrite its spacetime volume integral as a surface integral at the boundaries of space and time. While we don't concern ourselves with the details of Dirchlet vs. Neumann boundary conditions, we assume that there is some BC which makes the surface terms vanish:

$$\frac{\partial}{\partial t}\left(\delta\psi_\sigma \frac{\partial\mathcal{L}}{\partial\dot\psi_\sigma} \right) = \boldsymbol{\nabla}\cdot\left(\delta\psi_\sigma \frac{\partial\mathcal{L}}{\partial\dot\psi_\sigma} \right) = 0 \tag{8.11}$$

So, if we set $\delta S = 0$, this tells us that the first quantity in brackets must vanish if $\delta\psi_\sigma(\mathbf{x},t)$ has an arbitrary profile; we thus get the Euler-Lagrange equations:

$$\frac{\partial\mathcal{L}}{\partial\psi_\sigma(\mathbf{x},t)} - \frac{\partial}{\partial t}\frac{\partial\mathcal{L}}{\partial\dot\psi_\sigma(\mathbf{x},t)} - \boldsymbol{\nabla}\cdot\frac{\partial\mathcal{L}}{\partial\boldsymbol{\nabla}\psi_\sigma(\mathbf{x},t)} = 0$$

$$\frac{\partial\mathcal{L}}{\partial\psi^{\dagger\sigma}(\mathbf{x},t)} - \frac{\partial}{\partial t}\frac{\partial\mathcal{L}}{\partial\dot\psi^{\dagger\sigma}(\mathbf{x},t)} - \boldsymbol{\nabla}\cdot\frac{\partial\mathcal{L}}{\partial\boldsymbol{\nabla}\psi^{\dagger\sigma}(\mathbf{x},t)} = 0 \tag{8.12}$$

In "normal" classical mechanics, we don't have the third term, but in our case we do have a spatial dependence. A note: one can see that this already lifts nicely to relativistic physics. Another note: potentials can cause some complication, here.

We can get the field equations from the EL-equations, and we have already derived the commutator. There is still the issue of operator ordering. In this non-relativistic theory, it is the only place that we see it. But in general this is something one must keep in mind.

Other than recovering what we previously had, we actually get more! This is in the context of classical field theory, and it has to do with symmetries.

8.3 Symmetry and Noether's Theorem

The kinds of symmetry we are interested in are where there is some infinitesimal transformation (i.e. continuous symmetries). For example a rotation, or translation (an example of one without an infinitesimal transformation would be parity).

We take our field variable, and we transform it to some other field variable:

$$\psi_\sigma(\mathbf{x}, t) \to \tilde{\psi}_\sigma(\mathbf{x}, t) = \psi_\sigma(\mathbf{x}, t) + \delta\psi_\sigma(\mathbf{x}, t)$$
$$\psi^{\dagger\sigma}(\mathbf{x}, t) \to \tilde{\psi}^{\dagger\sigma}(\mathbf{x}, t) = \psi^{\dagger\sigma}(\mathbf{x}, t) + \delta\psi^{\dagger\sigma}(\mathbf{x}, t) \tag{8.13}$$

Where we can take the transformation to be linear as we consider infinitesimal transformations (higher orders negligeble). We can now consider how the Lagrangian transforms under this. We say that this is a *symmetry* (this is our definition) if $\delta\mathcal{L}$ can be organized (algebraically; without looking at the equations of motion) in the following way:

$$\delta\mathcal{L} = \frac{\partial}{\partial t}R(\mathbf{x}, t) + \boldsymbol{\nabla} \cdot \mathbf{J}(\mathbf{x}, t) \tag{8.14}$$

If things drop off at infinity, then this is a way of saying that the action doesn't change. Some textbooks require the Lagrangian to be invariant; we do not enforce this constraint here.

Now, having defined a symmetry, we can invoke the equations of motion and see what more we can learn. If we assume that the equation of motion is obeyed, then we have that:

$$\delta\mathcal{L} = \frac{\partial}{\partial t}\left(\delta\psi_\sigma\frac{\partial\mathcal{L}}{\partial\dot{\psi}_\sigma}\right) + \boldsymbol{\nabla} \cdot \left(\delta\psi_\sigma\frac{\partial\mathcal{L}}{\partial\boldsymbol{\nabla}\psi_\sigma}\right) + (\psi_\sigma \iff \psi^{\dagger\sigma}) \tag{8.15}$$

So we now have two equations for $\delta\mathcal{L}$; their difference should then be zero:

$$0 = \frac{\partial}{\partial t}\left[\delta\psi_\sigma(\vec{x}, t)\frac{\partial\mathcal{L}}{\partial\dot{\psi}_\sigma(\mathbf{x}, t)} + \delta\psi^{\dagger\sigma}(\vec{x}, t)\frac{\partial\mathcal{L}}{\partial\dot{\psi}^{\dagger\sigma}(\mathbf{x}, t)} - R(\mathbf{x}, t)\right]$$
$$+ \boldsymbol{\nabla} \cdot \left[\delta\psi_\sigma(\vec{x}, t)\frac{\partial\mathcal{L}}{\partial\boldsymbol{\nabla}\psi_\sigma(\mathbf{x}, t)} + \delta\psi^{\dagger\sigma}(\vec{x}, t)\frac{\partial\mathcal{L}}{\partial\boldsymbol{\nabla}\psi^{\dagger\sigma}(\mathbf{x}, t)} - \mathbf{J}(\mathbf{x}, t)\right] \tag{8.16}$$
$$\implies \frac{\partial}{\partial t}\mathcal{R}(\mathbf{x}, t) + \boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{J}}(\mathbf{x}, t) = 0$$

in other words, we have a conservation law! The symmetry implies a conservation law, and a systematic way of finding the conserved quantity; this is the power of the Lagrangian. This is *Noether's Theorem*. $\mathcal{R}$ is the Noether charge density and $\boldsymbol{\mathcal{J}}$ is the Noether current density. If we now integrate, we find:

$$\frac{\partial}{\partial t}\int_\Omega d^3x\,\mathcal{R}(\mathbf{x}, t) = -\oiint_{\partial\Omega} d\mathbf{n} \cdot \boldsymbol{\mathcal{J}}(\mathbf{x}, t) \tag{8.17}$$

In otehr words the change in the Noether charge is just the Noether current that flows across the boundary. We can now lift this to quantum field theories, where we will find great use of these results. Next time we will explore various examples of symmetries, and see what Noether's theorem has to say about their associated conserved quantities.

9 Symmetries & Noether's Theorem

Last time we reverted from quantum to classical field theory. We did this to cast our field theory in terms of a Lagrangian and equations of motion derived from them; this had the benefit of allowing us to use Noether's theorem, which tells us about conserved quantities and how to determine what they are. We can then lift these equations of motion back into our QFT, showing that operators in our QFT are also conserved. There are two problems in the way; the first is operator ordering issues, and the second is that products of operators at the same point requires some mathematical definition, which can be add odds with the conservation law! This can actually lead to interesting physics though; in particle physics these are known as anomalies, e.g. in pion decay. The miracle is that this fooling around with operators leads to effects that are experimentally observable. Abstract and seemingly ad-hoc mathematics have physical impact!

But we won't go there (at least for now); we will first study some examples of symmetries and Noether's Theorem. We need an example theory before we find symmetries of said theory; let's write down our favourite one:

$$\mathcal{L} = \frac{i\hbar}{2}\psi^\dagger\frac{\partial\psi}{\partial t} - \frac{i\hbar}{2}\frac{\partial\psi^\dagger}{\partial t}\psi - \frac{\hbar^2}{2m}\boldsymbol{\nabla}\psi^\dagger\cdot\boldsymbol{\nabla}\psi - \frac{\lambda}{2}(\psi^\dagger\psi)^2 \tag{9.1}$$

where we have omitted the **x**s and spin indices for brevity. The symmetrized time derivative part ensures that it is real. We also don't include chemical potential for now. But note it is useful for operator ordering issues; changing the order of terms generates extra constants, and the chemical potential can be used to absorb these ambiguities.

9.1 Phase Symmetry

One transformation that leaves $\mathcal{L}$ invariant is to multiply ψ by some phase (we can see this as each ψ is paired with $\psi^\dagger$). So we can write down an infinitesmal version of this transformation:

$$\begin{aligned}\delta\psi &= i\theta\psi \\ \delta\psi^\dagger &= -i\theta\psi^\dagger\end{aligned} \tag{9.2}$$

we can test that this is indeed a symmetry by seeing if we can write the Lagrangian as a total derivative under it; but it is even easier than that here; just plugging it in, to linear order we see that:

$$\delta\mathcal{L} = 0 \tag{9.3}$$

i.e. the Lagrangian density doesn't change at all. We call this *phase symmetry*. Going back to our expression from last class of the Noether current and Noether charge density, we have:

$$\mathcal{R}(\mathbf{x},t) = \delta\psi\frac{\partial\mathcal{L}}{\partial\dot{\psi}} + \delta\psi^\dagger\frac{\partial\mathcal{L}}{\partial\dot{\psi}} = \hbar\theta\psi^\dagger\psi = 0 \tag{9.4}$$

We see what we would have called the number operator in the QFT popping out on the RHS; in other words, the particle number is a conserved Noether charge of some kind. We can also write:

$$\mathcal{J} = \delta\psi\frac{\partial\mathcal{L}}{\partial\boldsymbol{\nabla}\psi} + \delta\psi^\dagger\frac{\partial\mathcal{L}}{\partial(\boldsymbol{\nabla}\psi^\dagger)} = 0. \tag{9.5}$$

so we also get information about the flux of the particle number. We have the following expression for the particle current:

$$\mathcal{J} = i\theta\frac{\hbar^2}{2m}(\boldsymbol{\nabla}\psi^\dagger)\psi - \frac{i\theta\hbar^2}{2m}(\psi^\dagger\boldsymbol{\nabla}\psi) \tag{9.6}$$

If this follows a continuity equation, then so will the above multiplied by a constant; so we learn:

$$\frac{\partial}{\partial t}(\psi^\dagger \psi) + \boldsymbol{\nabla} \cdot \left(\frac{i\hbar}{2m} \boldsymbol{\nabla} \psi^\dagger \psi - \frac{i\hbar}{2m} \psi^\dagger \boldsymbol{\nabla} \psi \right) = 0 \tag{9.7}$$

We can define the density:

$$\rho(\mathbf{x}, t) = \psi^{\dagger\sigma}(\mathbf{x}, t)\psi_\sigma(\mathbf{x}, t) \tag{9.8}$$

and the current:

$$\mathbf{J}(\mathbf{x}, t) = -\frac{i\hbar}{2m} \left(\psi^{\dagger\sigma}(\mathbf{x}, t)\boldsymbol{\nabla}\psi_\sigma(\mathbf{x}, t) - \boldsymbol{\nabla}\psi^{\dagger\sigma}(\mathbf{x}, t)\psi_\sigma(\mathbf{x}, t) \right) \tag{9.9}$$

So this is our first example of a conservation law (note: this follows from Noether's Theorem, which we proved last time!); phase symmetry leads to the conservation of particle number. Note that the above example is operator ordering ambiguous, so it should lift nicely to QFT; it tells us if our theory has a phase symmetry, then:

$$\frac{\partial}{\partial t}\mathcal{N} = \frac{\partial}{\partial t} \int d^3x \rho(\mathbf{x}, t) = 0 \tag{9.10}$$

We knew this already for specific examples, but now we have a more general argument; we know it to be true for any theory with a phase symmetry. Now, what else can we do?

9.2 Translational Symmetry and the Stress Tensor

We would expect that our theory is translation invariant in both time and in space, as there are now explicit ts or xs anywhere. These should lead to conservation of energy and momentum, respectively. A time translation would be:

$$\psi(\mathbf{x}, t) \to \psi(\mathbf{x}, t + \epsilon) \approx \psi(\mathbf{x}, t) + \epsilon \frac{\partial}{\partial t}\psi(\mathbf{x}, t) \tag{9.11}$$

and a space translation would be:

$$\psi(\mathbf{x}, t) \to \psi(\mathbf{x} + \boldsymbol{\epsilon}, t) \approx \psi(\mathbf{x}, t) + \boldsymbol{\epsilon}\boldsymbol{\nabla}\psi(\mathbf{x}, t) \tag{9.12}$$

So we could then write:

$$\delta\psi = (\epsilon\frac{\partial}{\partial t} + \boldsymbol{\epsilon} \cdot \boldsymbol{\nabla})\psi$$
$$\delta\psi^\dagger = (\epsilon\frac{\partial}{\partial t} + \boldsymbol{\epsilon} \cdot \boldsymbol{\nabla})\psi^\dagger \tag{9.13}$$

since $\mathcal{L}$ does not depend on t or $\mathbf{x}$, if we make the above infinitesimal transformations, the change in the Lagrangian would look like:

$$\delta\mathcal{L} = (\epsilon\frac{\partial}{\partial t} + \boldsymbol{\epsilon} \cdot \boldsymbol{\nabla})\mathcal{L} = \frac{\partial}{\partial t}(\epsilon\mathbf{L}) + \boldsymbol{\nabla} \cdot (\boldsymbol{\epsilon}\mathcal{L}) \tag{9.14}$$

so it satisfies a symmetry criterion, and therefore there is a conserved charge; in fact there are four, one for each component of ϵ (like a conserved four-vector)! Then there are 3 conserved currents for each of these. The best way to write this is as a 4x4 matrix; let's set this up.

$$\mathcal{R}(\mathbf{x}, t) = \left(\epsilon\frac{\partial}{\partial t} + \boldsymbol{\epsilon} \cdot \boldsymbol{\nabla} \right) \psi\frac{\partial\mathcal{L}}{\partial\dot{\psi}} + \left(\epsilon\frac{\partial}{\partial t} + \boldsymbol{\epsilon} \cdot \boldsymbol{\nabla} \right) \psi^\dagger\frac{\partial\mathcal{L}}{\partial(\boldsymbol{\nabla}\psi^\dagger)} - (\epsilon\mathcal{L}) \tag{9.15}$$

and there are one of these for each of the four ϵs. The current densities are analogous:

$$\mathcal{J}(\mathbf{x}, t) = \left(\epsilon\frac{\partial}{\partial t} + \boldsymbol{\epsilon} \cdot \boldsymbol{\nabla} \right) \psi\frac{\partial\mathcal{L}}{\partial(\boldsymbol{\nabla}\psi)} + \left(\epsilon\frac{\partial}{\partial t} + \boldsymbol{\epsilon} \cdot \boldsymbol{\nabla} \right) \psi^\dagger\frac{\partial\mathcal{L}}{\partial(\boldsymbol{\nabla}\psi^\dagger)} - \epsilon\mathcal{L} \tag{9.16}$$

we now write this as a matrix, known as the energy-momentum (stress) tensor:

$$\begin{pmatrix} \mathbb{T}^{tt} & \mathbb{T}^{ta} \\ \mathbb{T}^{bt} & \mathbb{T}^{ba} \end{pmatrix} \tag{9.17}$$

where the first row is the time row, and the latter three are the space rows (and analogously for the columns). a, b range over xyz. The conservation laws for each of these quantities looks like:

$$\frac{\partial}{\partial t} \mathbb{T}^{tt} + \frac{\partial}{\partial (x^b)} \mathbb{T}^{bt} = 0 \tag{9.18}$$

$$\frac{\partial}{\partial t} \mathbb{T}^{ta} + \frac{\partial}{\partial (x^b)} \mathbb{T}^{ba} = 0. \tag{9.19}$$

where the derivatives act on the first indices, and the second label points to the symmetry we are talking about. We can then read off the components of the stress tensor:

$$\mathbb{T}^{tt} = \frac{\partial \psi}{\partial t} \frac{\partial \mathcal{L}}{\partial \dot{\psi}} + \frac{\partial \psi^\dagger}{\partial (\dot{\psi}^\dagger)} \frac{\partial \mathcal{L}}{\partial (\dot{\psi}^\dagger)} - \mathcal{L} = \mathcal{H} = \frac{\hbar^2}{2m} \boldsymbol{\nabla} \psi^\dagger \cdot \boldsymbol{\nabla} \psi + \frac{\lambda}{2} (\psi^\dagger \psi)^2 \tag{9.20}$$

we recognize this as the Hamiltonian density. So what is conserved here is the integral of this over space, i.e. the total energy! The other components are:

$$\mathbb{T}^{ta} = \frac{\partial \psi}{\partial (x^a)} \frac{\partial \mathcal{L}}{\partial \dot{\psi}} + \frac{\partial \psi^\dagger}{\partial (x^a)} \frac{\partial \mathcal{L}}{\partial (\dot{\psi}^\dagger)} \tag{9.21}$$

note we've stripped off the ϵ as it is just an overall factor. We also don't include the $-\mathcal{L}$ term in the above. This we can recognize as the momentum density, whose integral is the spatial momentum; which is conserved. You can get the three space components by varying $a = x, y, z$, and each of these are conserved.

Note that writing the stress tensor does not do anything except save us some writing. In conclusion, we have the conserved quantities:

$$U = \int d^3 x \mathbb{T}^{tt}(\mathbf{x}, t) \tag{9.22}$$

$$P^a = \int d^3 x \mathbb{T}^{ta}(\mathbf{x}, t) \tag{9.23}$$

Does this lift to the quantum field theory? The momentum is easy because it turns out to be quadratic. Just looking at $\mathbb{T}^{ta}$, we can see that it only depends on the first terms of the Hamiltonian; it doesn't care about the interactions. Any interaction that has space translation invariance will have the same expression for the momentum density and the moementum itself. Moreover, it is the same as the number operator case where there is no operator ordering ambiguity.

However, there is an ambiguity for the $\mathbb{T}^{tt}$/energy term. This is more or less an ad-hoc procedure where we guess ordering until it works. Luckily we don't encounter this issue very much, and in this case we already know what the ordering is. So we know how to deal with it already.

9.3 Galilean Relativity

A teaser for next day: translation symmetry in space and time give us a tensor. There are some more symmetries, though; in a sense symmetries of a generic theory (symmetries that lots of theories share), which go in the direction of relativity. Even in non-relativistic physics, there are concepts of relativity. One example is Galilean symmetry/relativity (experiments in a moving car should yield the same results as that done in a stationary lab). On a primitive level, we can go back to Newtonian mechanics:

$$m\ddot{\mathbf{x}} = \mathbf{0} \tag{9.24}$$

and this equation has a Galilean symmetry. If we replace $\mathbf{x}(t)$ with $\tilde{\mathbf{x}}(t) = \mathbf{x}(t) + \mathbf{v}t$, we should still agree that Newton's second law holds. And they do, because the second derivative kills the $\mathbf{v}t$ term. How does one lift this to quantum mechanics? Well, we have a Schrodinger equation:

$$\left(i\hbar \frac{\partial}{\partial t} + \frac{\hbar^2 \boldsymbol{\nabla}^2}{2m} \right) \psi(\mathbf{x}, t) = 0 \tag{9.25}$$

How do we do a Galilean transformation here? Inside the wavefunction, let's try changing our coordinate; $\psi(\mathbf{x}, t) \to \psi(\mathbf{x} + \mathbf{v}t, t)$. But then this doesn't obey the original SE... we have to modify it:

$$\left(i\hbar \frac{\partial}{\partial t} - i\mathbf{v} \cdot \boldsymbol{\nabla} + \frac{\hbar^2 \boldsymbol{\nabla}^2}{2m} \right) \psi(\mathbf{x} + \mathbf{v}t, t) = 0. \tag{9.26}$$

This looks strange (we would want the new wavefunction to obey the same equation, after all). So what if we introduce a phase factor, i.e. consider:

$$\psi(\mathbf{x}, t) \to e^{i \frac{m\mathbf{v} \cdot \mathbf{x}}{\hbar}} \psi(\mathbf{x} + \mathbf{v}t, t) \tag{9.27}$$

But then we have:

$$\left(i\hbar \frac{\partial}{\partial t} + \frac{\hbar^2 \boldsymbol{\nabla}^2}{2m} + \frac{m\mathbf{v}^2}{2} \right) e^{i \frac{m\mathbf{v} \cdot \mathbf{x}}{\hbar}} \psi(\mathbf{x} + \mathbf{v}t, t) = 0 \tag{9.28}$$

(look in the notes for the correct formula...) so still our transformed wavefunction doesn't obey the SE. But it does look a bit closer, at the very least. So we have another redefinition:

$$\left(i\hbar \frac{\partial}{\partial t} + \frac{\hbar^2 \boldsymbol{\nabla}^2}{2m} \right) e^{i \frac{m\mathbf{v} \cdot \mathbf{x}}{\hbar} - i \frac{m\mathbf{v}^2}{2} t} \psi(\mathbf{x} + \mathbf{v}t, t) = 0. \tag{9.29}$$

There isn't a great intuition for why this looks so complicated... but in any case, this is it, and we can lift this to our quantum field equation (as part of our field equation looks exactly like the SE), and then for the interaction term there are no explicit $\mathbf{x}$s or ts and the whole theory is phase invariant, so these things cancel. Next day we will write down the infinitesimal transformations for this case, show that $\delta\mathcal{L}$ obeys the required relation, and write down a conservation law (but the quantity that is conserved is not so clear in this case; it will be derived quantity).

10 Galilean Symmetry

Last time we ended up the middle of a discussion of Galilean symmetry. We observed there that if $\psi(\mathbf{x}, t)$ was a solution of the Schrodinger equation, we could write down another solution:

$$e^{-\frac{i}{\hbar}\frac{mv^2}{2}t - i\frac{m}{\hbar}\mathbf{v}\cdot\mathbf{x}}\psi(\mathbf{x} + \mathbf{v}t, t). \tag{10.1}$$

This is a Galilean boost. Galilean symmetry is present in many classical mechanics systems. Landau and Lifschiz mention it in their textbook, where they argue that Galilean symmetry is the reason for why which $T = \frac{1}{2}mv^2$ (rather than any other power of v; see https://physics.stackexchange.com/questions/535/why-does-kinetic-energy-increase-quadratically-not-linearly-with-speed/14752#14752 for a very nice argument that uses Galilean symmetry to show this). This seems to be a symmetry of our equation of motion; let us now test if it is a symmetry of our full theory. We showed that it was a symmetry for the free theory; and it should also be a symmetry for our interacting theory of the type we have been discussing. If we either have hyper-local interactions or interactions that only depend on relative distance, they should have this symmetry.

10.1 Verifying the Symmetry

To test this on the level of Lagrangians, we first write down the infinitesimal version of the transformation. We let the velocity be infinitesimal and taylor expand to first order, and then strip off $\mathbf{v}$:

$$\delta\psi = t\boldsymbol{\nabla}^a\psi(\mathbf{x}, t) - i\frac{m}{\hbar}x^a\psi(\mathbf{x}, t) \tag{10.2}$$

And the $\psi^\dagger$ version of this is just replacing everything with its complex conjugate:

$$\delta\psi^\dagger = t\boldsymbol{\nabla}^a\psi^\dagger(\mathbf{x}, t) + i\frac{m}{2}x^a\psi^\dagger(\mathbf{x}, t) \tag{10.3}$$

To confirm that this is a symmetry, we calculate the first variation of the Lagrangian density, and use algebra to see if we can express it as in Eq. (8.14). In fact we can:

$$\delta\mathcal{L} = \boldsymbol{\nabla}^a(t\mathcal{L}(\mathbf{x}, t)) \tag{10.4}$$

Let us go through this calculation for the free-field Lagrangian, looking at it term-by-term.

$$\begin{aligned}
\delta(i\hbar\psi^\dagger\dot{\psi}) &= i\hbar\delta\psi^\dagger\dot{\psi} + i\hbar\psi^\dagger\frac{\partial}{\partial t}(\delta\psi) \\
&= i\hbar(t\boldsymbol{\nabla}^a\psi^\dagger + i\frac{m}{2x}x^a\psi^\dagger)\dot{\psi} + i\hbar\psi^\dagger(t\boldsymbol{\nabla}^a - i\frac{m}{2}x^a)\psi + i\hbar\psi^\dagger\boldsymbol{\nabla}^a\psi \\
&= i\hbar t\boldsymbol{\nabla}^a(\psi^\dagger\dot{\psi}) + i\hbar\psi^\dagger\boldsymbol{\nabla}^a\psi \\
&\quad + \boldsymbol{\nabla}^a(ti\hbar\psi^\dagger\dot{\psi}) + i\hbar\psi^\dagger\boldsymbol{\nabla}^a\psi
\end{aligned} \tag{10.5}$$

We have two terms; but the second term will be cancelled out by the next term in the Lagrangian density. It comes from the fact that $\delta\psi$ and $\delta\psi^\dagger$ have gradients $\boldsymbol{\nabla}^a$ acting on them and these do not commute with the x^as. So cancelling the extra bits, we end up with:

$$\delta\mathcal{L} = \boldsymbol{\nabla}^a(t\mathcal{L}(\mathbf{x}, t)) \tag{10.6}$$

We could also argue that the extra terms cancel out from the phase invariance. This demonstrates that this indeed is a symmetry from our technical definition. We could have expected this of course; we have a time dependent translation $\mathbf{x} \to \mathbf{x} + \mathbf{x}t$, and we had a phase factor out front to cancel out everything.

10.2 Noether's Theorem and Galilean Symmetry

The Noether charge density is given by:

$$B^a = t\mathbb{T}^{ta} + mx^a\rho \tag{10.7}$$

i.e. time times the noether charge density for spatial translation (momentum/the stress tensor!) plus the Noether charge for phase (density, where $\rho = \psi^\dagger\psi$). So this is the charge density conserved under Galilean boosts, and there are three of them corresponding to three directions. We also have a current density:

$$\mathcal{B}^{ab} = t\mathbb{T}^{ba} + mx^a J^b \tag{10.8}$$

again, these are just components of the stress tensor plus the particle current. We don't actually derive much new here. We recall that translation symmetry tells us:

$$\frac{\partial}{\partial t}\mathbb{T}^{ta} + \nabla_b\mathbb{T}^{ba} = 0 \tag{10.9}$$

and phase symmetry tells us:

$$\frac{\partial}{\partial t}\rho + \nabla_b J^b = 0 \tag{10.10}$$

And the above two together almost imply:

$$\frac{\partial}{\partial t}B^a + \nabla_b\mathcal{B}^{ab} = 0 \tag{10.11}$$

But not quite, because the $\frac{\partial}{\partial t}$ hits the t and the ∇ hits the x. We do have:

$$\mathbb{T}^{ta} + mJ^a = 0. \tag{10.12}$$

so this tells us that the momentum density is the mass times the particle current density; a very intuitive result, but one required by Galilean invariance. This is at the level of densities. What does it mean on a deeper sense? We can look at the conservation laws for Galilean symmetry. The following is a constant in time:

$$t\int d^3x\,\mathbb{T}^{ta} + \int d^3x\,mx^a\rho = \text{Constant} \tag{10.13}$$

where we have integrated the Noether charge density. We can think of the second term as the average position of the center of mass. Or, we can divide the entire equation by it and rewrite it as follows:

$$\int d^3x\,x^a\rho(\mathbf{x},t) = -\frac{1}{m}\left[\int d^3x\,\mathbb{T}^{ta}(\mathbf{x},t)\right]t + \langle x^a(0)\rangle \tag{10.14}$$

i.e. the average position of the COM is the initial position plus the average velocity times time... it simply means that the COM of the system moves at a constant speed. We already expected this, but now we have a concrete criteria to test for Galilean invariance!

10.3 Benefits of Lifting Galilean Symmetry to QFT

Can we learn something more practical from this? Let's look at something really cool; let's see what these symmetries tell us when we lift them to QFT. We won't go through the technicalities of this process (if it doesn't work, the problem tends to be really hard, or unsolvable anyway). We expect Galilean symmetry to survive this lifting[18] Let us see what the benefits of this symmetry are.

---

[18]Of course we are assuming that the system is non-relativistic; relativistic corrections are not Galilean invariant. In the relativistic case we would look at Lorentz invariance, instead.

Perhaps we want to calculate a correlation function like the following:

$$W(\mathbf{x_1}, t_1, \mathbf{x_2}, t_2) = \langle \mathcal{O} | \psi(\mathbf{x_1}, t_1) \psi^\dagger(\mathbf{x_2}, t_2) | \mathcal{O} \rangle. \tag{10.15}$$

Let's assume we have Galilean symmetry at the quantum level. Then presumably the ground state is Galilean invariant; even more primitively, it is translation invariant. Since it's an eigenstate of the Hamiltonian, time translation just gives rise to a phase which can be removed by shifting the Hamiltonian by a constant, anyway. Explicitly, we have the equation:

$$|\mathcal{O}\rangle = e^{-\frac{i}{\hbar}Ht}|\mathcal{O}\rangle \tag{10.16}$$

Space translation should look very similar, but instead of the Hamiltonian, we would put the momentum operator and space in the imaginary exponential:

$$|\mathcal{O}\rangle = e^{-\frac{i}{\hbar}\mathbf{P}\cdot\hat{a}}|\mathcal{O}\rangle \tag{10.17}$$

A homogenous system should have these symmetries. Such translations should be generated by unitary operations, for which the ground state is invariant under. For Galilean symmetry, the situation is more complicated, but we may think that there is similarly a unitary operator associated with a Galilean boost. Let us assume there is one without knowing its form; we know it should have the action:

$$|\mathcal{O}\rangle = U|\mathcal{O}\rangle \tag{10.18}$$

$$U^\dagger \psi(\mathbf{x}, t) U = e^{-\frac{i}{\hbar}\frac{mv^2}{2}t - i\frac{m}{\hbar}\mathbf{v}\cdot\mathbf{x}}\psi(\mathbf{x} + \mathbf{v}t, t) \tag{10.19}$$

in a system with this symmetry, U with these properties should exist. What do they tell us about the correlation functions? If we generate the time translation, and plug in the translated object into the correlation function (i.e. put in):

$$\langle \mathcal{O} | U^\dagger \psi(\mathbf{x_1}, t_1) U U^\dagger \psi^\dagger(\mathbf{x_2}, t_2) U | \mathcal{O} \rangle. \tag{10.20}$$

where U is the time translation, we learn that the correlation *is only a function of* $t_1 - t_2$. This is because the e^{iH} in the middle is only a function of $t_1 - t_2$ and the e^{iH} at the end doesn't depend on anything. The only dependence is on the difference! The exact same argument follows for positions, where one can prove the correlation only depends on $\mathbf{x_1} - \mathbf{x_2}$. In other words, symmetry has reduced the number of things that W can depend on from 8 to 4.

Rotation symmetry is not discussed yet, but it is in HW3! But it is really the same story. We will find that the ground state is rotation invariant, and so we can generate a rotation, and so we learn that W can only be a function of a rotationally invariant combination of $\mathbf{x_1}, \mathbf{x_2}$, i.e. W only depends on $(\mathbf{x_1} - \mathbf{x_2})^2, t_2 - t_1$ (only on the magnitude of the vector!!) So there are only two variables that the correlation functions can depend on! And we haven't even used Galilean symmetry yet! If we do apply it, we find[19]:

$$W((\mathbf{x_1} - \mathbf{x_2})^2, t_1 - t_2) = e^{\frac{im}{\hbar}\frac{(\mathbf{x_1} - \mathbf{x_2})^2}{(2(t_1 - t_2))}} f(t_1 - t_2) \tag{10.21}$$

so symmetry has reduced our problem appreciably. We can go no further with symmetry; for the free field theory f is a constant. Interactions seem to only modify this by multiplying by some function of time. We should feel powerful at this point; for Galilean symmetry we only need to find some f that depends on time. In conclusion: symmetry simplifies a lot! We can write objects such as correlation functions to depend on only one variable, for example.

---

[19]Note; there might be some singular behaviour if the positions/times are equal; some more work might be necessary in this case

10.4 Scale Invariance

The symmetries discussed so far have been in some sense "generic" in that they apply to many systems; but there are a few more interesting symmetries that occur only in very special systems. This perhaps makes them less interesting, but they have been studied in depth over the last 20 years or so, especially in systems such as cold atoms. These are symmetries relating to changing scales; scale transformations, and something related known as conformal transformations (also known as Special Schrodinger transformations). We will discuss them briefly here as they are likely to come up when one studies CM physics, cold atom physics etc. . .

If we put the interaction to zero and look at the free field equation, we have an equation with a scale symmetry:

$$\left(i\hbar \frac{\partial}{\partial t} + \frac{\hbar^2 \boldsymbol{\nabla}^2}{2m} \right) \psi(\mathbf{x}, t) = 0 \tag{10.22}$$

At first this might seem strange, as there appear to be dimensionful quantities like $\hbar, m$ in the above. But nevertheless, if $\psi(\mathbf{x}, t)$ is a solution, then so is

$$\Lambda^{d/2} \psi(\Lambda x, \Lambda^2 t) \tag{10.23}$$

where the Λ is the scaling factor, and the $\Lambda^{d/2}$ appears so that the transformed solution is still normalized. Going back to the example with the correlation functions, the scale invariance determines $f(t_1 - t_2) = (t_1 - t_2)^d$ times some overall constant and so up to a constant we have determined the correlation function completely! The two-point correlation functions have been completely solved for us. One might wonder; if we had interactions here, could we use this to solve interacting theories? The answer is yes, and this is what makes this very interesting. What makes this difficult is that if we add interactions, adding linear terms as operators in general messes up the scale symmetry; there are only few examples we can add. The Coulomb interaction is one that might seem scale-invariant (it doesn't seem to contain any dimensionful quantities at first) but when one tries to add it in, one finds that indeed scale invariance is violated. In the simple case we consider above, we are safe with no interactions. With interactions, if the coupling constant values are tuned very precisely (i.e. to "fixed points") then we do have scale invariance.

Note that since in the free field equation the equation is symmetric, we might expect the same of the action. Indeed, one can confirm this by looking at the infinitesimal transformation:

$$\delta \psi = \left(\mathbf{x} \cdot \boldsymbol{\nabla} - 2t \frac{\partial}{\partial t} + \frac{d}{2} \right) \psi \tag{10.24}$$

and if we construct the Noether charge and current density, because the above terms are like translations in space in time, algebraically what comes out is something that contains the stress tensor (again).

A philosophical question: Why is lifting classical field theory to QFT so successful? Because for many situations we have weak coupling (e.g. E&M, gravity, weak interactions) and so the structure is preserved when we take the classical behaviour to quantum (and of course this is why the classical approaches to these topics are quite accurate, and we study it in undergrad).

11 Scale and Conformal Symmetry

11.1 Non-Relativstic Conformal Field Theory

We've been studying the "space-time" symmetries of our non-relativistic QFT. Though space-time is a bit of a stretch, as the non-relativistic theories don't really have a grounding in geometry (though literature does exist on "non-relativistic" general relativity, casting Galilean transformations as isometries of non-relativistic spacetime. But this is a bit contrived...) In the real world, non-relativistic symmetries are just symmetries for things that are moving slowly (we should be able to derive them from SR by taking the non-relativistic limit)! This is a more useful point of view; in the world we live in, almost everything we would analyze with the kind of field theory we have been discussing is made of particles which have a relativistic foundation behind them. With that said, let us wrap up this discussion with a short introduction to non-relativistic CFT.

This isn't a particularly new subject; it's origin dates back to the 1970s. That said it has only been developed quite recently in the context of cold atoms, and critical behaviour (in the sense of a phase transition). The connection is that (in a rough sense); when something undergoes a phase transition (e.g. a paramagnet becoming ferromagnetic) there are fluctuations on all length scales; such that if we change the resolution at which we view the magnet, it does not look much different at different scales. This tends to characterize a phase transition; you get something roughly scale-invaraint. We started about the mathematical machinery behind scale invariance last day. If we have a theory that is symmetric under scale transformation, it looks the same under different scales!

We've written down a free field theory (so we actually have an example of a theory for which this holds) that has the property of scale invariance:

$$\left(i\hbar \frac{\partial}{\partial t} + \frac{\hbar^2 \boldsymbol{\nabla}^2}{2m} \right) \psi(\mathbf{x}, t) = 0 \tag{11.1}$$

so (for example) the degenerate Fermi gas we discussed has scale symmetry. A real Fermi gas may have interactions that break this scale invariance, but still perhaps in certain regimes it may look the same (e.g. a piece of copper looks the same $1\text{m}, 10\text{m}, 100\text{m}$ away; it is scale invariant at large distances/in the infrared, even though this may break at smaller scales). We identified the infinitesimal transformation:

$$\delta\psi = \left(\mathbf{x} \cdot \boldsymbol{\nabla} + 2t \frac{\mathrm{d}}{\mathrm{d}t} + \frac{d}{2} \right) \psi(\mathbf{x}, t) \tag{11.2}$$

and under this transformation, the Lagrangian density of this theory transforms (in what is known as the *scale transformation*) as:

$$\delta\mathcal{L} = \boldsymbol{\nabla} \cdot (\mathbf{x}\mathcal{L}) + \frac{\mathrm{d}}{\mathrm{d}t} (2t\mathcal{L}) \tag{11.3}$$

so it can be written in terms of total derivatives; it is therefore a symmetry! We can now turn the crank and write down the Noether charges and currents. But before that, we note that the above is accompanied often with a *conformal transformation* (also known as a Special Schrodinger transformation/symmetry), which reads:

$$\delta\psi = \left(t^2 \frac{\mathrm{d}}{\mathrm{d}t} + t\mathbf{x} \cdot \boldsymbol{\nabla} - i\frac{m}{\hbar} \frac{\mathbf{x}^2}{2} + \frac{d}{2}t \right) \psi \tag{11.4}$$

here it isn't really clear where this comes from; the origin will be clearer in the relativistic case. The free field Lagrangian density transforms as:

$$\delta\mathcal{L} = \frac{\partial}{\partial t} (t\mathcal{L}) + \boldsymbol{\nabla} \cdot (t\mathbf{x}\mathcal{L}) \tag{11.5}$$

so again this is a symmetry. We will see why conformal symmetry often comes along in a theory with scale symmetry (and we will see some similar things in the assignment, where the symmetry of the stress tensor

under the spatial indices implies rotational invariance. There is also a connection with translation/number invariance implying Galilean symmetry etc.)

11.2 Noether's Theorem for Scale Symmetry

We can write down the Noether charge as:

$$D = \delta\psi \frac{\partial \mathcal{L}}{\partial \dot{\psi}} + \delta\psi^\dagger \frac{\partial \mathcal{L}}{\partial(\dot{\psi}^\dagger)} - 2t\mathcal{L} \tag{11.6}$$

and the current as:

$$\boldsymbol{\mathcal{D}} = \delta\psi \frac{\partial \mathcal{L}}{\partial(\boldsymbol{\nabla}\psi)} + \delta\psi^\dagger \frac{\partial \mathcal{L}}{\partial(\boldsymbol{\nabla}\psi^\dagger)} - \mathbf{x}\mathcal{L} \tag{11.7}$$

A handy tip if we work with fermions to deal with the potential signs: If we have a Lagrangian, and we want to compute $\delta\psi \frac{\partial \mathcal{L}}{\partial \dot{\psi}}$; then it's second order in the ψs so it commutes with everything, until we get to ψ and we replace it. We can now go look up the Lagrangian and see what we get:

$$D = t2\mathbb{T}^{tt} + x_b \mathbb{T}^{tb} \tag{11.8}$$

The ath Noether charge is:

$$\boldsymbol{\mathcal{D}}^a = 2t\mathbb{T}^{at} + x_b \mathbb{T}^{ab} - \frac{d}{2}\frac{\hbar^2}{2m}\boldsymbol{\nabla}^a(\psi^\dagger\psi) \tag{11.9}$$

where we note the last $d/2$ piece gives us a bit of a mess (unlike the rest of the terms which come from the space/time transformations); it's like a phase transformation, but it's missing an i; so its really something else that we have not seen yet. This is indeed conserved:

$$\frac{\partial}{\partial t}D + \boldsymbol{\nabla}_a \boldsymbol{\mathcal{D}}^a = 0 \tag{11.10}$$

by Noether's theorem. If we assume translation invariance:

$$\partial_t \mathbb{T}^{tt} + \boldsymbol{\nabla}_a \mathbb{T}^{at} = 0$$
$$\partial_t \mathbb{T}^{tb} + \boldsymbol{\nabla}_a \mathbb{T}^{ab} = 0 \tag{11.11}$$

we then obtain:

$$2\mathbb{T}^{tt} + \sum_a \mathbb{T}^{aa} - \frac{d}{t}\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2\rho = 0. \tag{11.12}$$

where $\rho = \psi^\dagger\psi$. This almost looks like the trace of the stress tensor (up to the junk at the end, and the factor of 2 in the front), so this is *almost* saying that scale invariance implies a stress tensor.

11.3 Improving the Stress Tensor

The stress tensor has an ambiguity; in fact, any Noether current does. Way at the beginning when we test for a symmetry, we try to see if $\delta\mathcal{L}$ can be written as a total derivative using algebra. This is an ambiguous process if there exist quantities in the variation which can be added to the two derivative terms and cancel. This ambiguity then carries through to the whole calculation, to the point where there may be additional things you can add to the current... so then which current is which? There isn't a great answer for non-relativistic physics; in the relativistic case, we have gravity and so we can pick the current that couples to gravity. In NR we have no right to do this, but we can pick a different stress tensor that is still consistent with the physics. The improvement proceeds as follows. We add something to our previous stress tensor. If the previous stress tensor was symmetric it should be symmetric, and it should

43

obey the same conservation laws, and it should be written as some derivative. The candidate that satisfies all three of these qualifications are:

$$\tilde{\mathbb{T}}^{ab} = \mathbb{T}^{ab} - \frac{d}{d-1}\left(\delta^{ab}\boldsymbol{\nabla}^2 - \boldsymbol{\nabla}^a\boldsymbol{\nabla}^b\right)(\psi^\dagger\psi) \tag{11.13}$$

note that the time-time and time-space components to not change. With the above modification, note that we still have conservation, if $\mathbb{T}^{ab}$ is symmetric then so is $\tilde{\mathbb{T}}^{ab}$ etc. So indeed we have an ambiguity. Why is this improved? For this improved stress tensor, scale invariance implies:

$$2\tilde{\mathbb{T}}^{tt} + \sum_a \tilde{\mathbb{T}}^{aa} = 0 \tag{11.14}$$

We consider a homogenous/isotropic system, and take the expectation value of the above. The expectation value of the first term is the energy density, and the diagonal components of the stress tensor are pressures, which are all the same as we assume isotropy. We therefore find:

$$2U - dP = 0. \tag{11.15}$$

In other words:

$$U = \frac{d}{2}P \tag{11.16}$$

which is a prediction of scale invariance! We obtain an equation of state of a scale invariant theory. We computed this above for the degenerate Fermi gas; we can now go back and check that this indeed holds for $d = 3$! Also, note that Eq. (11.14) can now be used as a criterion for scale invariance. Also also; note that if we add interaction, all bets are off!

   Note: the SE may look like a scale invariant problem, but the SE always has a bound state with a binding energy; how can we have a binding energy for a theory with no dimensionful paramaters? The problem is we have a singular potential (e.g. a dirac delta) so we need some extra boundary conditions for the behaviour of the wavefunction at these points. This turns out to be equivalent to specifying the binding energy, but quantifying this violates scale invaraince. Finding scale invariant theories is hard. Though, one can usually tune parameters/coupling constants to get there; unfortunately this usually appears in places where calculations are intractable; either coupling constants are too large to use perturbation theory, or we set them all to zero and we have a trivial/Gaussian fixed point). Other fixed points have been shown to exist, e.g. the unitarity point where a state just gains some binding symmetry, or the point where phase symmetry gets broken etc.

11.4   Noether's Theorem for "Conformal" Symmetry

There's certainly less intuititon for the conformal symmetry... it's not really a conformal symmetry in that conformal symmetries preserve angles. The transformation in Eq. (11.4) doesn't really satisfy this. But, it is of interest, so we can write down the Noether current density (and let's assume we've done the whole song and dance of improving the stress tensor so as to get rid of the junk terms...) - actually we did not have time to do this. But let's see how we can get conformal symmetry:

   Let's assume we have translation symmetry in time and space as in Eq. (11.11) and also have phase symmetry:

$$\frac{\partial\rho}{\partial t} + \boldsymbol{\nabla}\cdot\mathbf{J} = 0 \tag{11.17}$$

and Galilean symmetry:

$$\mathbb{T}^{ta} = -mJ^a \tag{11.18}$$

and scale invariance:

$$2\tilde{\mathbb{T}}^{tt} + \sum_a \tilde{\mathbb{T}}^{aa} = 0 \tag{11.19}$$

These symmetries imply the conformal symmetry (intriguingly, we don't need rotational symmetry!).

Note that the effects of these symmetries constrain various (2/3) point correlation functions E.g. deuteron has a low binding energy relative to the masses, almost a critical system with conformal symmetry, one can do calculations of matrix elements and compare this to experimental scattering data etc.

Next class: So far, there haven't been unintuitive leaps to what we have done. Everything has been methodical; just doing mathematics (doing nothing)! We've just been playing with things; there haven't been challenges to ou; physical intuition. To go to the next level, we need to understand special relativity to some level, and we will construct these relativistic systems. We will do this by couching SR as a type of symmetry (a true symmetry of spacetime). We can then build field theories that have these symmetries. We have the technical tools at our disposal, but we do require a bit of a leap. What is different? What we've learnt about QM is still QM, we've just placed it in another context. But for example take the pi-meson generator at TRIUMF. To describe this, we go from a state with a few nuclei to a state with a whole bunch of pions. We need a QM theory where the number of particles is not fixed. QFT is a natural place for this. $E = mc^2$ tells you that if you have the right energies and quantum numbers, you can create particles. The FT description is ideal for this; we can write down dynamical processes that do these things. This is why QFT is often called the natural marriage of QM and special relativity. But there is another deep reason, namely causality. What happens is if we ask "what if we did single particle QM?" with position $\mathbf{x}$, momentum $\mathbf{p}$, with commutation relations $[x_i, p_j] = i\hbar\delta_{ij}$, and the relativistic Hamiltonian:

$$H = \sqrt{\mathbf{p}^2 c^2 + m^2 c^4}. \tag{11.20}$$

Things go very wrong! When we calculate wavepacket spreads, there is a nonzero probability that the particle could have travelled superliminally (because the wavepacket is a superposition of a huge range of possible momenta). We can show that sometimes we will detect it outside of the light cone. We violate some very fundamental principles. Causality is not a philosophical principle, but an experimental fact that we do not see particles that go faster than the speed of light[20]. So we would discard this theory because it doesn't describe nature (that said if we do find tachyons, maybe we dig this back up).

---

[20]Throwback to the experiment that detected "superliminal neutrinos" because of a loose cabel

12 Relativistic Quantum Mechanics, Space-Time Coordinates

12.1 The Naive setup

Let us discuss for a few minutes a subject that "doesn't exist" - we will see why through the discussion! Think of a quantum-mechanical particle with position $\mathbf{X}$, momentum $\mathbf{P}$, such that these follow the canonical commutation relations $[X_i, P_j] = i\hbar\delta_{ij}$. We will give this particle a relativistic expression for the energy:

$$H = \sqrt{P^2c^2 + m^2c^4} \tag{12.1}$$

We can now ask about this system. If we were naively trying to combine QM with SR, this would be the first kind of thing we would try. In this kind of system, we can discuss eigenstates of position and momentum (although of course these are not eigenstates in the true sense; they are not normalized):

$$H|p\rangle = \sqrt{p^2c^2 + m^2c^4}|p\rangle \tag{12.2}$$

the question of how this eigenstate evolves in time is a simple one; its a stationary state of the system, so only evolves according to some phase:

$$|p(t)\rangle = e^{-\frac{i}{\hbar}Ht}|p\rangle = e^{-\frac{i}{\hbar}\sqrt{p^2c^2+m^2c^4}t}|p\rangle. \tag{12.3}$$

Of course nothing interesting happens to eigenstates of the Hamiltonian with time. The more interesting question is what would happen to an eigenstate of position? We can characterize this as the matrix element $brax_f e^{-\frac{i}{\hbar}Ht}|x_i\rangle$ which is the probability (transition) amplitude that a particle starts at position x_i and some time later is found at position x_f. From the above, we can calculate this as:

$$\langle x_f|e^{-\frac{i}{\hbar}Ht}|x_i\rangle = \int d^3p\langle x_f|p\rangle e^{-\frac{i}{\hbar}\sqrt{p^2c^2+m^2c^4}t}\langle p|x_i\rangle \tag{12.4}$$

Where we have inserted the resolution of the identity $\int d^3p|p\rangle\langle p|$. The inner products of momentum and position we are familiar as:

$$\langle p|x\rangle = \frac{e^{-i\frac{\mathbf{p}}{\hbar}\cdot\mathbf{x}}}{(2\pi\hbar)^{3/2}} \tag{12.5}$$

So then:

$$\langle x_f|e^{-\frac{i}{\hbar}Ht}|x_i\rangle = \int \frac{d^3p}{(2\pi\hbar)^3}e^{i\frac{\mathbf{p}}{\hbar}\cdot(\mathbf{x}_f-\mathbf{x}_i)-\frac{i}{\hbar}\sqrt{p^2c^2+m^2c^4}t} \tag{12.6}$$

but this final expression we derive with the standard (naive) assumptions has problems.

12.2 Problem 1 - Lack of Lorentz Invariance

For one, it is not Lorentz invariant. Why should an probability distribution change if we put the experiment on a train? Or if we put ourselves on a train and interpret the experiment? It transforms something like a charge density; the time component of a four-vector. This is because we have setup $\psi^\dagger\psi$ to transform like the time-component of a four-vector, even though we want to normalize the wavefunction and this normalization should hold between all frames.

We could fix this up to be Lorentz invariant via an easy fix up; we insert a factor of the energy downstairs:

$$\langle x_f|e^{-\frac{i}{\hbar}Ht}|x_i\rangle = \int \frac{d^3p}{(2\pi\hbar)^3}\frac{mc^2}{\sqrt{pc^2+m^2c^4}}e^{i\frac{\mathbf{p}}{\hbar}\cdot(\mathbf{x}_f-\mathbf{x}_i)-\frac{i}{\hbar}\sqrt{p^2c^2+m^2c^4}t} \tag{12.7}$$

there is nothing in our theory that tells us why we do this, but it does work (and it will be something we get from QFT). When $pc^2 \ll m^2c^4$ (i.e. in the non-relativistic limit) the factor is one so we recover

the NRQM result. Another note: inserting this factor has significant effects, where we have difficulty localizing the particle (this carries through to quantum field theory - we cannot localize a particle to a position smaller than its wavelength. But in NRQM we can do this; the wavefunction of a particle can be as sharply peaked as we need it to be).

12.3 Problem 2 - Causality Violations
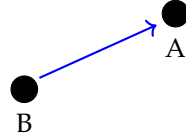
The next problem is harder to patch up.

Figure 12.1: Consider an experiment where Bob releases a particle and Alice is some distance away. Under naive relativistic QM, Alice has a non-zero chance of measuring the particle to be at her position, even if she is so far away such that light would not be able to reach her in that time from when Bob released it.

When we do the calculation for the probability of finding the particle some arbitrarily far distance away, we find that we have a non-zero probability regardless of the distance. To see this, let us return to the above transition probability. We can evaluate the integral approximately using the edge-of-the-wedge theorem. We note that $\langle x_f | e^{-\frac{i}{\hbar}Ht} | x_i \rangle$ is analytic in the lower half of the complex time plane; this comes from the fact that that there the $e^{-\frac{i}{\hbar}\sqrt{p^2c^2 + m^2c^4}t}$ leads to an exponential supression of the integral. As time comes to the real axis, it becomes much less nice (it becomes something like a distribution). We can define the value on the real axis as the limit taken from the analytic lower half time plane. How does this tell us anything? If the theory were causal, for small t we would have to find that the integral is zero. But then we have some arc in the complex plane for which the function is zero. But if this region is in the region where the function is analytic, then by analytic continuation we could show that it would be zero everywhere.

So the last step to show is that it is indeed analytic in this region. In the lower-t plane, we can use Cauchy's integral formula to write:

$$f(t_0) = \oint \frac{dt}{2\pi i} \frac{f(t)}{t - t_0} \tag{12.8}$$

where $f(t) = \langle x_f | e^{-\frac{i}{\hbar}Ht} | x_i \rangle$ is our integral of interest. The value of the integral does not depend on the contour taken; let us deform the contour so that part of the contour lies on the real axis. We can then push the t_0 up until it is on the real axis; at this point the formula still holds, and so $f(t_0)$ is an analytic function on the real axis as well. Therefore by analytic continuation, the integral is zero everywhere. Contradiction!

We therefore conclude that our theory violates causality. This would be ok if the universe actually behaved like this, but as far as we know we have never observed this; physics appears to be causal. So, this theory doesn't reflect reality, and we should throw it away.
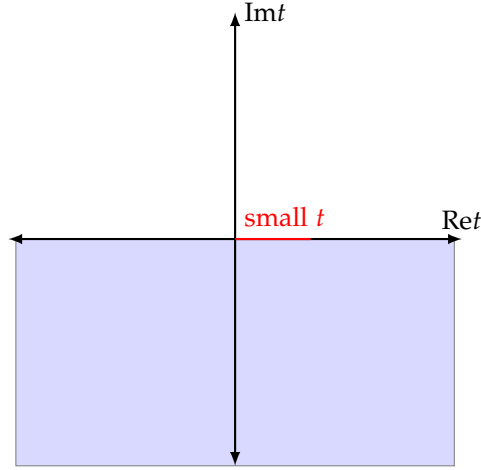
Figure 12.2: View of the complex t-plane. The integral for $\langle x_f | e^{-\frac{i}{\hbar}Ht} | x_i \rangle$ (Eq. (12.7)) is analytic in the lower-half plane due to the exponential supression in the integral. For the theory to be causal, the integral must vanish for small real t (in red). However, if this is indeed the case that the integral vanishes for an arc in $\mathbb{C}$ and is analytic on the arc, we may use analytic continuation (taylor expanding the integral along the arc where it vanishes) to conclude that the integral vanishes throughout the entire complex plane; a contradiction.
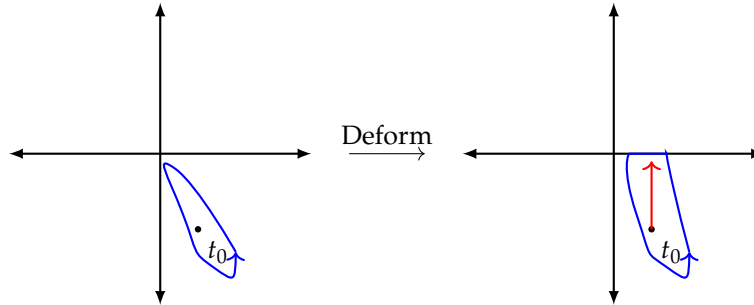

Figure 12.3: In the region where it is analytic (i.e. the lower-half plane), we can take $\langle x_f | e^{-\frac{i}{\hbar}Ht} | x_i \rangle$ at some t_0 to be defined by an arbitrary contour integral in the lower-half plane by Cauchy's Integral Formula. However, we can then deform the contour such that that part of the contour lies on the small real t region where the integral must vanish for the theory to be causal. We can then push up t_0 (take a limit) until it lies on the real axis in the small t region. But then Cauchy's integral formula still holds there, so in fact the integral is analytic on the real axis, and in particular is analytic and vanishing in the small t region; it therefore vanishes everywhere by the previous analytic continuation observation (giving us the desired contradiction).

12.4 The QFT fix-up

How does QFT fix this? First, we get the factor of $\frac{mc^2}{\sqrt{pc^2+m^2c^4}}$ to preserve lorentz invariance. But the other thing we do is we make use of the negative energy branch of the energy spectrum:

$$E = -\sqrt{p^2c^2 + m^2c^4}. \tag{12.9}$$

When we calculate the same process in QFT, something gets added! We have two amplitudes:

$$\langle x_f|e^{-\frac{i}{\hbar}Ht}|x_i\rangle = \int \frac{d^3p}{(2\pi\hbar)^3}\left[\frac{mc^2}{\sqrt{pc^2+m^2c^4}}e^{i\frac{\mathbf{p}}{\hbar}\cdot(\mathbf{x}_f-\mathbf{x}_i)-\frac{i}{\hbar}\sqrt{p^2c^2+m^2c^4}t} - \frac{mc^2}{\sqrt{pc^2+m^2c^4}}e^{i\frac{\mathbf{p}}{\hbar}\cdot(\mathbf{x}_f-\mathbf{x}_i)+\frac{i}{\hbar}\sqrt{p^2c^2+m^2c^4}t}\right] \tag{12.10}$$

i.e. we add a term that looks like time is running backwards. How do we intereprt this? The first term/process has the interpretation of Bob having a particle and Alice detects it. The second term/process is as follows. When trying to detect a particle, Alice produces a particle-antiparticle pair, the antiparticle which goes the other direction and annhilates Bob's particle. We find in QFT that outside of the light cone that these two processes perfectly destructively interfere, and so causality is preserved!

An interesting note: if we repeated this analysis with a massless particle, we would find that the probability only has support on the lightcone (As we would expect; massless particles should travel on the lightcone!)

Another note: the addition of this second term destroys our analyticity argument as we now have one term that is analytic in the lower-t plane and the other is analytic in the upper-t plane. So we do not get the contradiction that the integral vanishes everywhere, as we did before.

Does this fixup fix everything? It seems to. They seem to be Lorentz invariant, and causal. This is why QFT gains the title of the natural marriage of quantum mechanics and special relativity (and string theory is known as the natural marriage of QM and GR - each are the only currently known solution to unifying these theories). Although this cannot be said of string theory yet, QFT at least seems to pass some very stringent experimental tests.

12.5 Special Relativity and Minkowski Spacetime

Once we go to a relativistic setting, it is natural to discuss space-time transformations (which are still translations, rotations and the like) which are symmetries of the space-time in which the QFT lives. We then try to write down the QFT such that these symmetries are not violated. We already know how to deal with this a little bit; we know Noether's theorem, and how to turn the crank and find conservation laws. We will therefore proceed to approach symmetries from a different point of view; as arising from the geometries of spacetime.

The space-time we consider is Minkowski spacetime; as much as we love curved spacetime, we won't consider it in this course here. It is described by coordinates:

$$\mathbf{x} = (x^1, x^2, x^3) \equiv x^a, \quad a = 1, 2, 3 \tag{12.11}$$

which describes three-dimensional infinite Euclidean space. The indices on the xs now have a position that means something. It is always up on a coordinate (or at least for the next while - x_a is meaningless, for now). We will add to this another coordinate for time to make this a space-time (but we will give it units of distance to make it symmetric):

$$x^0 \equiv ct \tag{12.12}$$

where x^0 ranges over the real line. Combining these, we obtain a four-vector:

$$x^\mu = (x^0, x^1, x^2, x^3), \quad \mu = 0, 1, 2, 3. \tag{12.13}$$

We will generally use latin letters for regular vectors, and greek letters for four-vectors (though sometimes we leave it off entirely when the object is clear from context). The cartesian coordinates are not the only coordinates possible, but they are quite convenient; symmetries are quite apparent here.

12.6 Coordinates on a space-time

If we have a general spacetime, we could think about a coordinate system; all this would be is a dictionary for the various points in the spacetime, with points x^μ, y^μ etc. There is of course a bunch of mathematical structure behind this, e.g. continuity/smoothness, this is done formally by locally mapping points (homeomorphism) to $\mathbb{R}^n$. We won't need this formality, but we will takeaway the idea of a coordinate transformation. Suppose we have a different coordinate system, with points $\tilde{x}^\mu, \tilde{y}^\mu$ etc. We could then compare dictionaries and figure out how to translate in between them. We could then define a transformation in between them: $\tilde{x}^\mu(x^\mu)$. This is a mathematical way of writing down a transformation, and the transformation inherits the "smoothness" of the coordinates (it is differentiable, invertible...). Formally, the invertiblity criterion is:

$$\det(\frac{\partial}{\partial x^\mu}\tilde{x}^\nu) \neq 0. \tag{12.14}$$

12.7 Scalar Fields

We consider a scalar field $\phi(x^\mu)$; this is a physical entity that takes a value at each point in spacetime (for example; the temperature in a room). If someone else comes along with a different set of coordinates, we should agree with the value of the scalar field; so we should agree that:

$$\phi(x) = \tilde{\phi}(\tilde{x}) \tag{12.15}$$

so a scalar field is defined in this way; with how it undergoes coordinate transformations.

We will discuss next time, but we can consider that displacement transforms like:

$$dx^\nu = \frac{\partial x^\mu}{\partial \tilde{x}^\nu}d\tilde{x}^\nu \tag{12.16}$$

So perhaps a vector field transforms in the same way:

$$A^\mu(x) = \frac{\partial x^\mu}{\partial \tilde{x}^\nu}d\tilde{A}^\nu(x) \tag{12.17}$$

13 Fields, Metrics, and Space-Time Symmetry

13.1 Review - Coordinates, Transformations, Scalar Fields

Last time, we were considering a coordinate system for our spacetime (which has requirements such as smoothness that we take as a given); we have a four-vector x^μ which represents a point in the spacetime. We then considered a different coordinate system $\tilde{x}^\mu$ and a change-of-coordinates transformation $x^\mu \to \tilde{x}^\mu(x)$ (four smooth functions in four variables). This function should be invertible, with $\tilde{x}^\mu \to x^\mu(\tilde{x})$ the inverse.

We can then discuss things that live in our spacetime, such as the scalar field. The term scalar says something about this; namely it tells us how we should translate values of the field under different coordinate transformations. We have some scalar field $\phi(x)$, and in some different set of coordinates we may have $\tilde{\phi}(\tilde{x})$ but at the same point the data should match, so:

$$\phi(x) = \tilde{\phi}(\tilde{x}) \tag{13.1}$$

13.2 Vector Fields

This is a straightforward generalization of a vector field; at each point in space we associate a vector with magnitude and direction. If temperature was an example of a scalar field, then wind would be a good example of a vector field; the wind has some magnitude and direction at each point in space. But now something to consider; when we transform between coordinates, we not only have to translate magnitudes, but also the direction of the wind. We say a (contravariant) vector field should transform like a differential:

$$dx^\mu = \frac{\partial x^\mu}{\partial \tilde{x}^\mu} d\tilde{x}^\nu \tag{13.2}$$

so:

$$A^\mu(x) = \frac{\partial x^\mu}{\partial \tilde{x}^\nu} \tilde{A}^\nu(\tilde{x}) \tag{13.3}$$

note the up and down indices here are quite important. x always has an up index. The index on the derivative is a bit interesting. We consider a transformation of the derivative:

$$\frac{\partial}{\partial x^\nu} = \frac{\partial \tilde{x}^\nu}{\partial x^\mu} \frac{\partial}{\partial \tilde{x}^\nu} \tag{13.4}$$

and we see that this is upside-down compared to the vector field transformation; so something different is happening here. We could find a vector field that transforms under this "upside-down" rule, and we would distinguish this by putting a down index on it (this would be a covariant) vector field:

$$A_\mu(x) = \frac{\partial \tilde{x}^\nu}{\partial x^\mu} \tilde{A}_\nu(\tilde{x}) \tag{13.5}$$

Note: we remind the reader that we work with Einstein summation convention, and so when we have two indices paired they are summed over. Since the partial derivative transforms covariantly, it is natural to think of it as an object as a down index:

$$\partial_\mu = \frac{\partial}{\partial x^\mu}, \quad \tilde{\partial}_\mu = \frac{\partial}{\partial \tilde{x}^\mu} \tag{13.6}$$

13.3 Tensor Fields

Tensor fields are a generalization of vector fields; they are things with more than one index. One example is the inertia tensor with two indices takes the form of a 3x3 matrix in 3-dimensional space. The relativistic description of an EM field is another. We can lift our transformations of vectors into tensors in the most obvious way:

$$\mathbb{T}^{\mu_1\cdots\mu_m}{}_{\nu_1\dots n}(x) = \frac{\partial x^{\mu_1}}{\partial \tilde{x}^{\rho_1}} \cdots \frac{\partial x^{\mu_m}}{\partial \tilde{x}^{\rho_m}} \frac{\partial \tilde{x}^{\sigma_1}}{\partial x^{\nu_1}} \cdots \frac{\partial \tilde{x}^{\sigma_n}}{\partial x^{\nu_n}} \tilde{\mathbb{T}}^{\rho_1\cdots\rho_m}_{\sigma_1\dots\sigma_n}(\tilde{x}) \tag{13.7}$$

13.4 Proper Time and Spacetime Metrics

At this point, we know nothing about the spacetime beyond the fact that I know I can index it and measure things. We need more than that; we will need to embed objects into this spacetime somehow. We will need some rules for doing so, and rules thar we can all agree on.

Let's take a pet snail[21] and put him at x^μ in our spacetime. Now let's say he moves to $x^\mu + dx^\mu$. We then record the time that elapses on the wristwatch. This we would call the proper time.
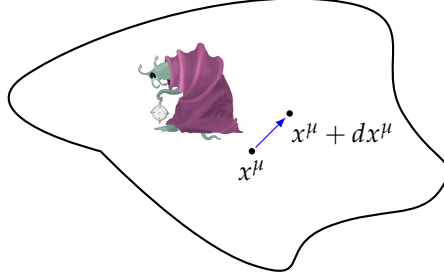

Figure 13.1: We take our pet snail (actually pictured is a slug, but close enough) and put him at coordinate x^μ in our spacetime. We then define the proper time as the time that elapses on his watch when he moves from x^μ to $x^\mu + dx^\mu$.

This should vary linearly with how far he travells. We should also have a perscription for the path taken to be somehow "sensible"; say $x^\mu + t dx^\mu$ for $0 \le t \le 1$. We want some way of figuring out what this proper time is. If it's linear in x^μ, but this has an index and proper time is just a number. So it would be useful to square the propert time; this would be bi-linear in x^μ, perhaps with an expression:

$$d\tau^2 = dx^\mu g_{\mu\nu}(x) dx^\nu \tag{13.8}$$

this defines a matrix $g_{\mu\nu}$ of some kind. It turns out to be a fairly special one; why? If we both watch the snail and record the proper time in the same way, we should agree on that proper time for every possible path he should take. We can therefore map out this matrix. Even if we restrict ourselves to timelike paths, there is enough to map out this tensor. The physics tells us that the times must agree, and so:

$$d\tau^2 = dx^\mu g_{\mu\nu}(x) dx^\nu = d\tilde{x}^\mu \tilde{g}_{\mu\nu}(\tilde{x}) d\tilde{x}^\nu \tag{13.9}$$

This tells us that g transforms as a tensor field! And so:

$$g_{\mu\nu}(x) = \frac{\partial \tilde{x}^\rho}{\partial x^\mu} \frac{\partial \tilde{x}^\sigma}{\partial x^\nu} \tilde{g}_{\rho\sigma}(\tilde{x}). \tag{13.10}$$

We give this tensor field a name of being a *metric*, or *metric tensor*. This encodes a lot of information about our spacetime, as it in a sense measures distances. We will not go into the direction of curvature and general relativity here, but this is certainly a fascinating direction.

13.5 Minkowski Spacetime

We are only example in a specific example of spacetime - Minkowski spacetime. We define it as a spacetime where there exists a coordinate system such that:

$$g_{\mu\nu}(x) = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{13.11}$$

---

[21] Let's assume that snails exist...

the poor snail notices that there are limits on his motion coming from the minus sign; he cannot for example go backwards in time. Note we have to correct the formula in Eq. (13.9) we had for $d\tau^2$ due to this minus sign:

$$d\tau^2 = -dx^\mu g_{\mu\nu}(x)dx^\nu = -d\tilde{x}^\mu \tilde{g}_{\mu\nu}(\tilde{x})d\tilde{x}^\nu \tag{13.12}$$

Note that this is known as the east-coast convention; one could alternatively take:

$$g_{\mu\nu}(x) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{13.13}$$

and this would be the west-coast convention. Despite our location at UBC, we adopt the east-coast convention.

Some background; the WC convention was developed by high-energy physicists, while the EC convention was developed by relativists. The former has the benefit of energies being positive, the latter has the benefit of distances being positive. Both are correct and often one takes the convention most convenient for the setting (depending on whether you discuss particle energies, or geometry).

13.6   Infinitesimal Coordinate Transformations

We consider infinitesimal coordinate transformations:

$$\tilde{x}^\mu = x^\mu + f^\mu(x) \tag{13.14}$$

We can think of f^μ as small but otherwise arbitrary. It also looks like a vector field, but is not quite one (it coordinate transforms in a funny way). Scalar fields transform under infinitesimal transformations as:

$$\phi(x) = \tilde{\phi}(\tilde{x}) = \phi(x) + \delta\phi(x) + f^\mu(x)\partial_\mu\phi(x) \tag{13.15}$$

where we have Taylor expanded in f for the last term. As usual we consider only up to linear order. Cancelling terms, we obtain:

$$\delta\varphi(x) = -f^\mu(x)\partial_\mu\varphi(x) \tag{13.16}$$

We can then abstract this to various other fields. One example is the contravariant vector field:

$$A^\mu(x) = \frac{\partial x^\mu}{\partial \tilde{x}^\nu}\tilde{A}^\nu(\tilde{x}) \tag{13.17}$$

Note the beauty of infinitesimal transformations in that they are very easy to invert:

$$x^\mu = \tilde{x}^\mu - f^\mu(\tilde{x}) \tag{13.18}$$

and so

$$\frac{\partial x^\mu}{\partial \tilde{x}^\nu} = \delta^\mu_\nu - \partial_\nu f^\mu \tag{13.19}$$

therefore:

$$A^\mu(x) = A^\mu(x) - \partial_\nu f^\mu(x)A^\nu(x) + \delta A^\mu(x) + f^\nu(x)\partial_\nu A^\mu(x) \tag{13.20}$$

From here the zeroth order bit cancells out, and so:

$$\delta A^\mu(x) = -f^\nu\partial_\nu A^\mu(x) + \partial_\nu f^\mu(x)A^\nu(x) \tag{13.21}$$

where the first term comes from the transformation of the coordinate and the second term comes from the transformation of the frame. The scalar field has the first part (and in fact every field will have it)! The extra terms are the derivatives of the transformation function with the indices matched up. Doing the same for covariant fields:

$$\delta A_\mu(x) = -f^\nu(x)\partial_\nu A_\mu(x) - \partial_\mu f^\nu(x)A_\nu(x) \tag{13.22}$$

And as a particular example, the metric transforms as:

$$\delta g_{\mu\nu}(x) = -f^\lambda(x)\partial_\lambda g_{\mu\nu}(x) - \partial_\mu f^\lambda(x)g_{\lambda\nu}(x) - \partial_\nu f^\lambda(x)g_{\mu\lambda}(x) \tag{13.23}$$

13.7 The definition of space-time symmetry

Having discussed metrics, what else can one say about a space? One thing that would be interesting is symmetry. Maybe we can define symmetry of spacetime as follows; symmetry is a change of coordinates such that the rules for calculating proper time remain exactly the same. We need the metric to calculate the proper time; so if we calculate the same proper time under a coordinate transformation, the metric better be left identical. Putting this together, we can say that *A symmetry of space-time is a coordinate transformation such that* $\delta g_{\mu\nu}(x) = 0$. So to discover symmetries, we can dream up a large class of fs, and see for what class we have $\delta g_{\mu\nu}(x) = 0$. This wouldn't be a useful definition of symmetry was plentiful, and in general it is quite rare to find them. Precisely, the maximum number of spacetime symmetries turns out to be $d(d+1)/2$[22]. This is obtained by counting the solutions to the differential equation obtained by setting $\delta g_{\mu\nu}(x) = 0$. An equivalent definition of symmetry is therefore: $\hat{f}^{\mu}(x)$ *is a symmetry if:*

$$\hat{f}^{\lambda}(x)\partial_{\lambda}g_{\mu\nu}(x) + \partial_{\mu}\hat{f}^{\lambda}(x)g_{\lambda\nu}(x) + \partial_{\nu}\hat{f}^{\lambda}(x)g_{\mu\lambda}(x) \tag{13.24}$$

i.e. $\hat{f}^{\mu}(x)$ *the Killing vector satisfies the Killing equation.* One might wonder if the form of the equation changes in different coordinate systems; this turns out to not be the case (and would require some differential geometry knowledge). This gives us a systematic way to discuss symmetries; something we will continue with next class.

---

[22]The number of independent components of a $d \times d$ symmetric tensor - though typically there are less, there are only 3 maximally symmetric spaces, namely Minkowski, and the spaces of constant positive or constant negative curvature. For 2D surfaces, this means the only maximally symmetric surfaces are the plane, sphere, and the one that looks locally like a Pringle (I missed the formal name that Gordon said here). In 4D we have Minkowski, DeSitter and Anti DeSitter space.

14 Symmetries of Minkowski Space-Time, Conformal Transformations

14.1 Metric Tensors

We have been discussing the metric tensor field $g_{\mu\nu}(x)$. This is special because it is used to measure proper times. There is one small thing we have forgot to mention which we should now fill in. What is $g^{\mu\nu}(x)$? The way we define this is the inverse of the metric with down indices. Therefore:

$$g^{\mu\nu}(x)g_{\mu\lambda}(x) = \delta^\mu_\lambda \tag{14.1}$$

w here δ^μ_λ is the 4×4 matrix with ones on the diagonal and zeros everywhere else. We can study how this transforms (and the answer is: it will not!)

$$\delta^\mu_\lambda = \frac{\partial x^\mu}{\partial \tilde{x}^\rho}\frac{\partial \tilde{x}^\sigma}{\partial x^\lambda}\tilde{\delta}^\sigma_\rho \tag{14.2}$$

the above equation only holds if $\tilde{\delta}^\sigma_\rho$ is identically a unit matrix. So, it looks the same in every coordinate system. It is sometimes called an *invariant tensor* for this reason (of which there are fairly few).

Note that $g^{\mu\nu}$ is also a tensor (a contravariant one).

Note that we have here assumed that the metric is invertible. If it is not, then our proper time experiments would have some strange results indeed... the snail could travel and no time could pass on the wristwatch, for example. This doesn't seem very physical. So we say that this shouldn't be possible; the metric should be invertible everywhere (technical terminology: the metric should be non-degenerate).

14.2 Contracting Fields, Raising/Lowering Indices

We can consider the contraction of two vector fields:

$$A_\mu(x)A'^\mu(x) \tag{14.3}$$

this object transforms like a scalar field; this makes sense, as it should transform like something without any indices.

In the same vein, once we have this tensor g, we can use it to raise indices on vector (tensor) fields:

$$A^\mu(x) = g^{\mu\nu}(x)A_\nu(x) \tag{14.4}$$

and similarly:

$$A_\mu(x) = g_{\mu\nu}A^\nu(x) \tag{14.5}$$

A note: haven't discussed spinors and their transformations; this is because there isn't really a transformation law for them. This might be a bit disquieting that this simple structure we have been exploring does not cover all possible physical objects (e.g. electrons are described by spinors). This will eventually be something we need to consider.

14.3 Symmetries of Space-Time

A symmetry of space-time is a coordinate transformation:

$$x^\mu \to \tilde{x}^\mu = x^\mu + f^\mu(x) \tag{14.6}$$

that leaves the metric invariant:

$$\delta g_{\mu\nu}(x) = 0. \tag{14.7}$$

So literally, g looks the same in both coordinate systems. We can't tell if we are doing physics in one coordinate systems or the other.

An example of a coordinate transformation that *isn't* a symmetry: the transformation that takes us from cartesian to polar coordinates. This is *not* a symmetry transformation. The schrodinger equation looks different in this coordinate system!

An example of a coordinate transformation that *is* a symmetry: a transformation that rotates/changes the direction of the Cartesian axes. In this case the laws of physics look exactly the same.

We can write explicitly the Killing equation:

$$\hat{f}^\lambda(x)\partial_\lambda g_{\mu\nu}(x) + \partial_\mu \hat{f}^\lambda(x)g_{\lambda\nu}(x) + \partial_\nu \hat{f}^\lambda(x)g_{\mu\lambda}(x) = 0. \tag{14.8}$$

The solutions $\hat{f}^\lambda(x)$ of which are called Killing vectors.

14.4 Killing Equation and Vectors for Minkowski Space-Time

Let's look at the Killing vectors for Minkowski space. Recall that Minkowski space is defined by the fact that we can go to a coordinate system where the metric looks like:

$$g_{\mu\nu}(x) = \eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{14.9}$$

We can plug this into the equation and find the Killing equation for Minkowski space:

$$\partial_\mu f^\lambda \eta_{\lambda\mu} + \partial_\nu f^\lambda \eta_{\mu\lambda} = 0. \tag{14.10}$$

If we now define:

$$\hat{f}_\mu(x) = \eta_{\mu\nu}\hat{f}^\nu(x) \tag{14.11}$$

then the Killing equation reads:

$$\partial_\mu \hat{f}_\nu + \eta_\nu \hat{f}_\mu = 0. \tag{14.12}$$

and we can start searching for our Killing vectors. Note that the above is a homogenous equation, which helps us with our search (we can look for solutions of zeroth order and linear order). They are:

1. $\hat{f}_\mu = C_\mu$, constants (of which there are 4). These correspond to time and space translation symmetry.

2. $\hat{f}_\mu = \omega_{\mu\nu}x^\nu$. We observe that this is indeed a solution with $\omega_{\mu\nu}$ anti-symmetric, so $\omega_{\mu\nu} = -\omega_{\nu\mu}$. These are the Lorentz transformations (in the broader sense); 3 spatial rotations and 3 Lorentz boosts.

Higher order polynomials do not solve the equation; this is it! This is as many Killing vectors as a space-time can have (Minkowski space is maximally symmetric) - there are 10 total for $D = 4$.

14.5 Conformal Transformations

Under a conformal transformation, the rules change slightly (this is a slight extension of our concept of symmetry). This is important as there are some QFTs that have this kind of enhanced symmetry (not technically a symmetry of the space as we have defined it). Conformal transformations are coordinate transformations for which the metric transforms to something proportional to the metric:

$$\delta g_{\mu\nu}(x) = \Omega(x)g_{\mu\nu}(x). \tag{14.13}$$

This is the Conformal Killing equation. In Minkowski space, it looks like:

$$\partial_\mu \hat{f}_\nu(x) + \partial_\nu \hat{f}_\mu(x) - \frac{2}{D}\eta_{\mu\nu}\partial_\lambda \hat{f}^\lambda(x) = 0. \tag{14.14}$$

Where D is the dimension of the space. One can check that the Killing vectors for Minkowski space are divergence free, so the third term vanishes and (of course) we have that the original Killing equation is satisfied. How many additional solutions are there? Not too many. Again it is homogenous so we can try the same strategy of looking for polynomial solutions. We find the conformal Killing vectors:

1. $\hat{f}_\mu C_\mu$

2. $\hat{f}_\mu = \omega_{\mu\nu} x^\nu$ for ω antisymmetric.

3. $\hat{f}^\mu = \lambda x^\mu$. A scale transformation.

4. $\hat{f}^\mu = x^\mu x^\nu b_\nu - b^\mu \frac{x^\nu x_\nu}{2}$ where b is some constant four-vector (of which there are four possible choices).

In the above, we have x with a lower index. It is defined as:

$$x_\nu = \eta_{\mu\nu} x^\mu \tag{14.15}$$

once we are in Minkowski space and agree on using Cartesian coordinates, we can raise and lower the indices on x (and on anything) with η. We can think of this as an inner product/space-time invariant:

$$x^\nu x_\nu = \eta_{\mu\nu} x^\mu x^\nu = \mathbf{x}^2 - (x^0)^2 \tag{14.16}$$

and this may be familiar from a previous course in special relativity.

As discussed before, conformal invariance can give us some extra symmetries when studying some systems; e.g. the classical Ising model on its second order phase transition; this turns out to be a QFT with $\hbar = 1$! It's not a quantum system, yet it obeys the rules of a QFT with a conformal symmetry. The conformal symmetry and some of its constraints have been exploited to calculate the critical exponents of the Ising model to world-record accuracy, matching numerical simulations perfectly.

14.6   Finite Lorentz Transformations

So, we've discussed the infinitesimal transformations. Let's say a few words about the finite case. We consider the Lorentz transforms:

$$x^\mu \to \tilde{x}^\mu = x^\mu + \omega^\mu_\nu x^\nu \tag{14.17}$$

So we write this as:

$$\tilde{x}^\rho = \Lambda^\rho_\mu x^\mu \tag{14.18}$$

where:

$$\Lambda^\rho_\mu \approx \delta^\rho_\mu + \omega^\rho_\mu \tag{14.19}$$

We consider the corresponding transformation for the metric:

$$\eta_{\mu\nu} = \frac{\partial \tilde{x}^\rho}{\partial x^\mu} \frac{\partial \tilde{x}^\sigma}{\partial x^\mu} \eta_{\rho\sigma} = \Lambda^\rho_\mu \Lambda^\sigma_\nu \eta_{\rho\sigma} \tag{14.20}$$

Lorentz transformations are the set of all Λs which obey the above equation. We have been studying the infinitesimal ones. We know that:

$$(\det \Lambda)^2 = 1 \tag{14.21}$$

and so:

$$\det \Lambda = \pm 1 \tag{14.22}$$

The $\det \Lambda = 1$ transformations are known as proper. It can be shown that the Λs here form a group. If Λ_1, Λ_2 obey the equation, then so do Λ_1^{-1}, $\Lambda_1 \Lambda_2$ etc.

Some objects are invariant under Lorentz transformations; for example scalars, and by extension any contraction of vector fields $A^\mu(x) A_\mu(x)$.

15 Real Scalar Field Theory

We've been discussing special relativity; now we use it to build up our relativistic quantum field theory. When we discussed our non-relativistic QFT, we didn't discuss geometry all too much; but here it will be useful to inform us about the kinds of space-time symmetries that our theory should possess as well. Our world doesn't look fairly symmetric; that said physics has had quite a bit of success exploring theories with symmetries to explain our universe; and elementary particle physics does appear to possess these.

15.1 Setting up the Real Scalar Field Theory

Let's look at a simple (in fact the simplest) example of a relativistic QFT. This will be simple in the sense that we will not have to worry about many indices. We consider a real scalar field theory. Consider a scalar field $\phi(x)$, and we want to consider a dynamical system where ϕ is a degree of freedom, and is relativistic. We can start writing down the Lagrangian density guided by what we have learned in the last few lectures. Some are convention dependent; there could be derivative terms in ϕ, and these are usually written first. The normalization of the derivative term fixes the normalization of the field ϕ, as well. We want to organize the derivatives such that they have Lorentz symmetry. We thus obtain the term:

$$-\frac{1}{2}\partial_\mu\phi(x)\partial^\mu\phi(x) \tag{15.1}$$

We could have terms with more derivatives (so-called effective field theories); while they would not be excluded in classical field theories, quantum field theories do not like them very much; so we do not consider them. We also could have something quadratic in ϕ:

$$-\frac{m^2}{2}\phi^2(x) \tag{15.2}$$

We could in principle have a term linear in ϕ, but we can get rid of it assuming that the Lagrangian has a symmetry such that it is unchanged under replacing ϕ with $-\phi$.

There are other terms that could contain ϕ, such as:

$$-\frac{\lambda}{4!}\phi^4(x) \tag{15.3}$$

(of course all the odd terms have to vanish by the $-\phi \leftrightarrow \phi$ symmetry). Of course we could have in principle ϕ^6, ϕ^8 terms etc. but these are not allowed due to renormalizability conditions. So in principle we have written everything down that could be contained in our Lagrangian density:

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\phi(x)\partial^\mu\phi(x) - \frac{m^2}{2}\phi^2(x) - \frac{\lambda}{4!}\phi^4(x) \tag{15.4}$$

We have tried to organize the Lagrangian density such that it transforms like a scalar under Lorentz transformations. Let's begin by studying what this Lagrangian tells us. The Euler-Lagrange equation in relativistic form for a scalar field is very nice; it reads:

$$\partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi(x))} = \frac{\partial\mathcal{L}}{\partial\phi(x)} \tag{15.5}$$

We thus find the equation of motion:

$$(-\partial^2 + m^2)\phi(x) + \frac{\lambda}{3!}\phi^3(x) = 0. \tag{15.6}$$

For λ finite this is very hard to solve, so in general the approach is to treat this in perturbation theory with small λ. If we set $\lambda = 0$, we obtain the Klein-Gordon equation:

$$(-\partial^2 + m^2)\phi(x) = 0. \tag{15.7}$$

and this term in (15.6) is the Klein-Gordon term, with wave operator $-\partial^2 + m^2$. The $\frac{\lambda}{3!}\phi^3(x)$ non-linear term can be thought of the interaction term. This is in analogy to our previous non-relativistic QFT, where we had a free field equation and added on the interaction terms. Note there isn't really a place here for an operator ordering ambiguity; everything commutes with everything here.

15.2 Natural Units

We work in a system of units where $\hbar = 1, c = 1$. Note that if we had not made this choice, then there would be factors of $\hbar, c$ in the $-\frac{m^2}{2}\phi^2(x)$ term. We can of course figure out where these should go by studying the units, but with natural units we are saved the hassle.

15.3 Commutation Relations

Note that this is the kind of system where the Lagrangian has velocity squared times some function of position:

$$\mathcal{L} = \frac{1}{2}\frac{\partial\phi}{\partial t}\frac{\partial\phi}{\partial t} + \dots \tag{15.8}$$

so the canonical momentum is $P = \frac{\partial\phi}{\partial t}$, and the canonical position $X = \phi$. We then have the Poisson bracket:

$$\{\phi, \frac{\partial}{\partial t}\phi\} = \delta() \tag{15.9}$$

and lifting this to the QFT, we have the commutator relation:

$$[\phi(x), \frac{\partial}{\partial t}\phi] = i\hbar\delta() \tag{15.10}$$

Concretely, we have the equal-time commutation relations of:

$$[\phi(x), \frac{\partial}{\partial y^0}\phi(y)]\delta(x^0 - y^0) = i\delta(x - y) \tag{15.11}$$

Of course the positions commute with each other:

$$[\phi(x), \phi(y)]\delta(x^0 - y^0) = 0 \tag{15.12}$$

As do the momenta:

$$[\frac{\partial}{\partial x^0}\phi(x), \frac{\partial}{\partial y^0}\phi(y)]\delta(x^0 - y^0) = 0 \tag{15.13}$$

One might wonder if particle physicists are missing any theories by only considering Lagrangian ones; but this is a discussion for another day.

15.4 Symmetry

We recall the infinitesimal transformation of the scalar field:

$$\delta\phi(x) = -f^\mu(x)\partial_\mu\phi(x) \tag{15.14}$$

where if this is a space-time symmetry then f^μ is a Killing vector; otherwise it is a generic change of coordinates. Let us keep it general for now; we see from algebra that:

$$\delta\mathcal{L}(x) = -f^\mu(x)\partial_\mu\mathcal{L}(x) + (\partial_\mu f_\nu + \partial_\nu f_\mu)\frac{1}{2}\partial^\mu\phi\partial^\nu\phi \tag{15.15}$$

We need to show that this transforms like the derivative of something. So let us rearrange using the product rule:

$$\delta\mathcal{L}(x) = \partial_\mu(-f^\mu(x)\mathcal{L}(x)) + (\partial_\mu f_\nu + \partial_\nu f_\mu)\frac{1}{2}\left(\partial^\mu\phi\partial^\nu\phi + \eta^{\mu\nu}\mathcal{L}\right) \tag{15.16}$$

Now if f^μ is a Killing vector, then $\partial_\mu f_\nu + \partial_\nu f_\mu = 0$ from the Killing equation in Minkowski space! So $\delta\mathcal{L}$ can be written as a total derivative - it is a symmetry in the Noetherian sense if f^μ is a Killing vector/we have a space-time symmetry:

$$\delta\mathcal{L}(x) = -f^\mu(x)\partial_\mu\mathcal{L}(x) \tag{15.17}$$

Applying Noether's theorem, we can obtain the Noether charge density and current. By Noether's theorem:

$$\mathcal{J}^\mu = \delta\phi\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi} - \mathcal{R}^\mu = \hat{f}^\nu\partial_\nu\phi\partial^\nu\phi + \hat{f}^\mu\mathcal{L} \tag{15.18}$$

where $\mathcal{R}^\mu = (R, \mathbf{J})$ from our old formulation of Noether's theorem (not in relativistic notation). So by Noether's theorem, this is conserved:

$$\partial_\mu\mathcal{J}^\mu = 0. \tag{15.19}$$

15.5 A Faster Noether's Theorem

We came here with a variation of φ where f was arbitrary, obtaining:

$$\delta\mathcal{L}(x) = \partial_\mu(-f^\mu(x)\mathcal{L}(x)) + (\partial_\mu f_\nu + \partial_\nu f_\mu)\frac{1}{2}\left(\partial^\mu\phi\partial^\nu\phi + \eta^{\mu\nu}\mathcal{L}\right) \tag{15.20}$$

Now, what if we actually imposed the equation of motion (something we have been told not to do)? If we vary the Lagrangian density and impose the the equations of motion (from the EL equation), then the above variation in $\mathcal{L}(x)$ has to be written as a total derivative. How do we make th esecond term in the above expression a total derivative? By writing $\partial_\mu(\dots)$ and $\partial_\nu()$... but then this implies the $\partial_\mu\left(\partial^\mu\phi\partial^\nu\phi + \eta^{\mu\nu}\mathcal{L}\right) = 0$ which gives us the result in a quick way!

15.6 The Stress Tensor

Another connection; let us write the Noether current as a contraction of the Killing vector:

$$\mathcal{J}^\mu = \hat{f}^\nu\left(\partial_\nu\varphi\partial^\mu\varphi + \eta^\mu_\nu\mathcal{L}\right) \tag{15.21}$$

We may call the above expression in the brackets a stress tensor $\mathbb{T}^\mu_\nu(X)$ and from the conservation:

$$\partial_\mu\mathbb{T}^{\mu\nu}(x) = 0. \tag{15.22}$$

So, in our context here we have the Stress tensor definition:

$$\mathbb{T}^{\mu\nu}(x) = \partial_\nu\varphi\partial^\mu\varphi + \eta^\mu_\nu\mathcal{L} \tag{15.23}$$

So: the Nother current for a symmetry $\delta\varphi = -\hat{f}^\mu\partial_\mu\varphi$ is:

$$\mathcal{J}^\mu = \hat{f}_\nu(x)\mathbb{T}^{\mu\nu}(x). \tag{15.24}$$

and this conserved stress tensor summarizes the conservation laws of spacetime symmetries. There are no operator ordering ambiguities, and the theory is not constructed to have any other symmetries, so we are in some sense done. But we still learn things. The time-time component gives the energy density, the time-space component gives the momentum density, and the other components give us stresses/strains and so on. Recall we also had this in non-relativistic physics, and this is exactly the analog. The construction of things like angular momentum operators from $\mathbb{T}$ in the non-relativistic setting generalizes here, by considering the various space-time symmetries of the form:

$$\hat{f}_\nu(x) = C_\nu + \omega_{\nu\lambda}x^\lambda. \tag{15.25}$$

15.7 Non-Interacting Theory

In order to gain intuition, let us try to solve this theory in the limit where it is solvable, i.e. when we set the nonlinear term to $\lambda = 0$. In this case, the field equation looks as follows:

$$(-\partial^2 + m^2)\varphi(x) = 0 \tag{15.26}$$

This can be solved (as usual) by the plane wave ansatz:

$$\varphi(x) \sim e^{ik_\mu x^\mu} \tag{15.27}$$

this solves the equation if:

$$k^\mu k_\mu + m^2 = 0. \tag{15.28}$$

Which tells us the relativistic dispersion relation:

$$k^0 = \pm\sqrt{\mathbf{k}^2 + m^2} \tag{15.29}$$

We need both solutions. The plane waves are intrinsically complex, but of course φ is an intrinsically real scalar field, i.e. $\varphi = \varphi^*$ in classical field theory of $\varphi = \varphi^\dagger$ in quantum field theory. We need plane waves on both sides to superimpose them and cancel out the imaginary components:

$$\varphi(x) = \int \frac{d^3k}{\sqrt{(2\pi)^3 2\omega(k)}} \left(e^{i\mathbf{k}\cdot\mathbf{x} - i\omega t} a(\mathbf{k}) - e^{-i\mathbf{k}\cdot\mathbf{x} + i\omega t} a^\dagger(\mathbf{k}) \right) \tag{15.30}$$

where $\omega(k) = \sqrt{\mathbf{k}^2 + m^2}$. One could check that the plane waves were complete, orthonormal (and so on) and that this is a solution to the QFT. In the QFT the $a, a^\dagger$s become operators with commutation relations:

$$[a(\mathbf{k}), a^\dagger(\mathbf{l})] = \delta^3(\mathbf{k} - \mathbf{l}) \tag{15.31}$$

which implies that:

$$[\varphi(x), \frac{\partial}{\partial y^0}\varphi(y)]\delta(x^0 - y^0) = \delta^3(\mathbf{x} - \mathbf{y}) \tag{15.32}$$

In fact if we look at the Hamiltonian:

$$H = \int d^3x \, \mathbb{T}^{00}(x) \tag{15.33}$$

and we plug in our solution and do some algebra, we find:

$$H = \int d^3k \, \omega(\mathbf{k}) a^\dagger(\mathbf{k}) a(\mathbf{k}) + C \tag{15.34}$$

i.e. a bunch of (non-interacting) harmonic oscillators plus a constant (which comes from the operator ordering), which many QFT courses tell you is harmless (as we are only interested in relative energies) but is not quite physically true... of course we have things like gravity which depend on the total energy density. But since we don't feel ourselves gravitating to some infinite energy density from the scalar field, then we should set it to zero to be consistent with physics:

$$H = \int d^3k \, \omega(\mathbf{k}) a^\dagger(\mathbf{k}) a(\mathbf{k}). \tag{15.35}$$

This in some sense is an ambiguity of the theory (and one of the mysteries of QFT)! It is an arbitrary constant, but gets a bit hairy when you think about it too much; e.g. the dark energy density in our theory

here would be infinite, but it is finite in our universe. We won't need it for anything because we won't be venturing off into quantum gravity, but it would be important there.

The momentum are:

$$P^a = \int d^3x \mathbb{T}^{0a}(x) = \int d^3k \, k^a a^\dagger(\mathbf{k}) a(\mathbf{k}) + C \tag{15.36}$$

here we might be able to argue that the constant is zero by an isotropy argument, but we run into troubles with frame dependence etc.... but let us not worry about it too much and just set it to zero.

Finally, let us construct quantum states for our theory. We have the vacuum $|0\rangle$, which is normalized:

$$\langle 0|0\rangle = 1 \tag{15.37}$$

and is annhilated by the annhilation operators:

$$a(\mathbf{k})|0\rangle = 0, \quad \langle 0|a^\dagger(\mathbf{k}) = 0 \quad \forall \mathbf{k} \tag{15.38}$$

We then have our basis:

$$\left\{ |0\rangle, a^\dagger(\mathbf{k})|0\rangle, \frac{a^\dagger(\mathbf{k}_1)a^\dagger(\mathbf{k}_2)|0\rangle}{\sqrt{2}}, \ldots \right\} \tag{15.39}$$

and we now more or less have a complete solution of our theory. As soon as we turn on λ we have no solution, but can obtain a solution via perturbation theory, where we learn how to include corrections. The whole course here on in is learning how to include these corrections in a systematic way, and in fact this is all we know how to do (other than trying to solve the theory numerically). For this particular theory, things have been studied up to λ^8, with roughly 8! Feynman diagrams...