Notes On Advanced Quantum Field Theory
The Theory of Elementary Interactions
A Course Given By Dr. Tobias Osborne

Transcribed by Dr. Alexander V. St. John

May 17, 2019

Lecture 1: Introduction

The goal for this course is to explain the current "standard model" for particle physics. This is too lofty of a goal for this course, so what we focus on is the textitbuilding blocks of the standard model, such that we understand the origin and purpose of each term of the Lagrangian.

Topics covered include

1. Path integral quantization

   Via Gaussian integrals.

2. Review perturbation theory via path integrals

   Includes familiar tools for calculating correlation functions such as Wick's theorem, Feynman rules, etc.

3. Renormalization

   Allows us to discuss effective QFT (e.g., eliminating infinities) in more detail.

4. Abelian and nonabelian classical gauge theories

   Uses path integrals to deduce quantizations from classical field theories.

5. Quantization of non-abelian gauge theories

   Employs path integrals for perturbative calculations and lattices for nonperturbative calculations.

6. Spontaneous symmetry breaking mechanisms

Path Integrals

Let's first suppose that the quantization is already done, and we have a quantum system with a Hilbert space $\mathcal{H}$, a Hamiltonian $\hat{H}$, and a propagator from integrating the Schroedinger equation, $U(t) = e^{-i\hat{H}t}$.

Now work out a representation to first order for the propagator by Taylor expanding

$$U(t) = e^{-i\hat{H}t} = \left(e^{-\frac{it}{N}\hat{H}}\right)^N = \lim_{N\to\infty}\left(\mathbb{I} - \frac{it}{N}\hat{H}\right)^N \tag{1}$$

Let $\{|j\rangle\}$ be a basis for the Hilbert space $\mathcal{H}$ and consider the transition amplitude of evolving from an eigenstate $|\phi_i\rangle$ to another eigenstate $|\phi_f\rangle$

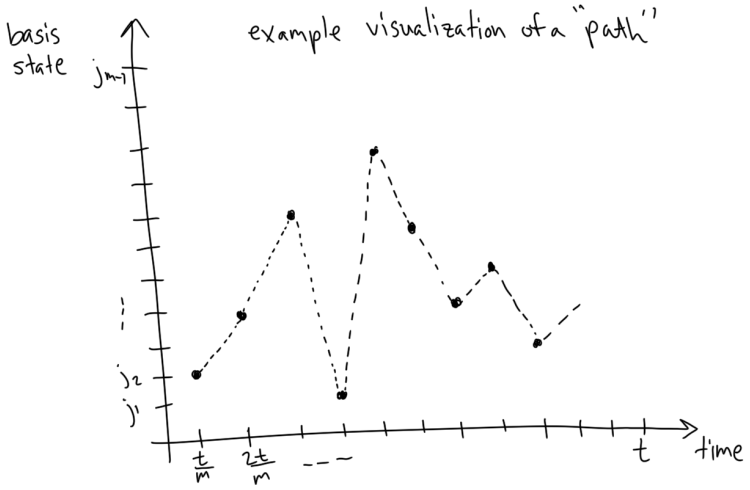
$$\langle\phi_f|\, U(t)\, |\phi_i\rangle = \langle\phi_f|\left(e^{-\frac{it}{N}\hat{H}}\right)^N |\phi_i\rangle \tag{2}$$

Insert $N-1$ Hilbert space basis completeness relations, one in between each of the N exponentials, and note that the sum over all paths from initial state to final state of functions of the paths

$$\langle \phi_f | \, U(t) \, | \phi_i \rangle = \sum_{j_1, \ldots, j_{N-1}} \langle \phi_f | \, e^{-\frac{it}{N}\hat{H}} \, | j_{N-1} \rangle \, \langle j_{N-1} | \, e^{-\frac{it}{N}\hat{H}} \ldots | j_1 \rangle \, \langle j_1 | \, e^{-\frac{it}{N}\hat{H}} \, | \phi_i \rangle \tag{3}$$

$$\equiv \sum_{\text{paths}} f(j_1, \ldots, j_{N-1}). \tag{4}$$

So, the transition amplitude of this state evolution is a sum over all of the paths through the basis states $j_1, \ldots, j_{N-1}$. An example schematic of a path is visualized below.



To work out the function of the path $f(j_1, \ldots, j_{N-1})$, we need to calculate these individual transition amplitudes between successive states $|j_{k-1}\rangle$ to some final state $\langle j_k|$ in the path integral setting, and seek to write it as an exponential of some function of the states $\mathcal{L}(j_{k-1}, j_k)$, the Lagrangian density

$$\langle j_k | \, (\mathbb{I} - \frac{it}{N}\hat{H}) \, | j_{k-1} \rangle \simeq e^{\frac{it}{N}\mathcal{L}(j_{k-1}, j_k)}. \tag{5}$$

So, the full transition amplitude will be a product of exponentials pf Lagrangian densities, well-known classical quantities

$$\langle \phi_f | \, U(t) \, | \phi_i \rangle = \sum_{\text{paths}} f(j_1, \ldots, j_{N-1}) = \sum_{j_1, \ldots, j_{N-1}} e^{\frac{it}{N} \sum_{k=2}^{N-1} \mathcal{L}(j_{k-1}, j_k)}. \tag{6}$$

Before moving forward, what makes the path integral so interesting?

1. It allows the calculation of quantum quantities, the transition amplitudes, via well-understood classical solutions and methods for handling highly oscillatory integrals such as the saddle point method.

2. It can also be used to build nonperturbative approximation schemes, such as Monte Carlo sampling over paths.

Example: General Nonrelativistic Quantum Mechanical System

Let's assume a little bit more about our quantum system. Suppose the quantum system is inspired by a classical system with pairs of canonical coordinates and momenta and the Hamiltonian $H(\{q^j\},\{p^j\}) = H(q,p)$.

Turn around the path integral sum over paths to guess a quantum Hamiltonian and Hilbert space from this classical Hamiltonian via

$$U(q_i, q_f, T) = \langle q_f | \, U(T) \, | q_i \rangle = \langle q_f | \, e^{-iT\hat{H}} \, | q_i \rangle \tag{7}$$

Proceed as before, mulitplying the exponentials and inserting $N-1$ completeness relations in between the N copies of the exponential. The completeness relation for the canonical position basis, a continuous variable, is

$$\mathbb{I} = \left(\prod_j \int dq_k^j \right) |q_k\rangle \langle q_k| \tag{8}$$

So, the transition amplitude in this case is, with $\epsilon \equiv \delta t = \frac{T}{N}$

$$\langle q_f | \, U(T) \, | q_i \rangle = \sum_{k_1,\ldots,k_{N-1}} \langle q_f | \, e^{-i\epsilon\hat{H}} \, | q_{k_{N-1}} \rangle \langle q_{k_{N-1}} | \ldots | q_{k_1} \rangle \langle q_{k_1} | \, e^{-i\epsilon\hat{H}} \, | q_i \rangle \, . \tag{9}$$

There are three cases for the dependence of the quantum Hamiltonian on the canonical coordinates in the expression for the propagator. It can depend purely on position, purely on momenta, or most realistically, it can depend on both.

In the case that the Hamiltonian is a function purely dependent on canonical position, such that $\hat{H} = g(\hat{q})$, we easily calculate the transition amplitude, which relates the quantum and classical canonical positions, since the $|q_k\rangle$ are energy eigenstates of the position-dependent Hamiltonian

$$\langle q_{k+1} | \, g(\hat{q}) \, | q_k \rangle = g(q_k) \prod_j \delta(q_k^j - q_{k+1}^j) \tag{10}$$

$$= g\left(\frac{q_{k+1} + q_k}{2} \right) \left(\prod_j \int \frac{dp_k^j}{2\pi} \right) e^{i \sum_j p_k^j (q_{k+1}^j - q_k^j)} . \tag{11}$$

Where we used the Dirac delta distribution identity $\delta(q) = \int_{-\infty}^{\infty} \frac{dp}{2\pi} e^{ip\cdot q}$ to introduce the canonical momenta into the transition amplitude. Also note that the Dirac delta function forces $q_{k+1} = q_k$, such that $f(q_k) = f(\frac{q_{k+1}+q_k}{2})$, and we write it in this fashion for later use.

Next, in the case that the Hamiltonian is a function purely dependent on canonical momenta, such that $\hat{H} = h(\hat{p})$, the transition amplitude is calculated by inserting the completeness relation for the momentum eigenbasis.

$$\langle q_{k+1}| \, h(\hat{p}) \, |q_k\rangle = \langle q_{k+1}| \, h(\hat{p}) \cdot \prod_j \int dp_k^j \, |p_k\rangle \, \langle p_k|q_k\rangle \tag{12}$$

$$= \prod_j \int \frac{dp_k^j}{2\pi} \, h(p_k) e^{i\sum_j p_k^j (q_{k+1}^j - q_k^j)} \tag{13}$$

Where the inner product of the position and momentum eigenstates is a Fourier phase element $\langle p|q\rangle = \frac{1}{2\pi} e^{ip\cdot q}$, and we get the sum, since the subscript k denotes N total canonical coordinate pairs.

The more realistic situation is when the Hamiltonian is dependent on both position and momenta $\hat{H} = \hat{H}(\hat{q}, \hat{p}) = g(\hat{q}) + h(\hat{p})$. Suppose the dependencies are linearly separable in the quantum Hamiltonian. Then we may translate between classical position and momenta via the Taylor expansion to first order

$$e^{-i\epsilon\hat{H}} = \mathbb{I} - i\epsilon\hat{H} = \mathbb{I} - i\epsilon(g(\hat{q}) + h(\hat{p})). \tag{14}$$

Using this linearity, we can write this dependence in the derived formula as

$$\langle q_{k+1}| \, e^{-i\epsilon\hat{H}(\hat{q},\hat{p})} \, |q_k\rangle = \prod_j \int \frac{dp_k^j}{2\pi} \, e^{-i\epsilon H(\frac{q_{k+1}+q_k}{2}, p_k)} e^{i\sum_j p_k^j (q_{k+1}^j - q_k^j)} \tag{15}$$

Putting all this together into the propagator, which is really the transition amplitude for a nonrelativistic quantum system,

$$U(q_i, q_f; T) = \left(\prod_{jk} \int dq_k^j \int \frac{p_k^j}{2\pi} \right) e^{i\sum_k \left(\sum_j p_k^j (q_{k+1}^j - q_k^j) - \epsilon H(\frac{q_{k+1}+q_k}{2}, p_k) \right)}. \tag{16}$$

Take note that there is nothing quantum on the RHS: no hats! We have used purely classical data to define the quantum propagator, or, transition amplitude, on the LHS, such that $U(q_i, q_f; T) \propto e^{-i\epsilon H(q_i, q_f; T)}$.

A few other remarkable points:
Using the saddle point method, we can build an approximation scheme for U. This is useful for solving highly oscillatory integrals, as we see in the transition

amplitude above (e.g., $e^{i\cdots}$), since such integrals can be approximated by its saddle points (or critical points), which correspond to classical paths of the system.

Monte Carlo sampling of the system can also be used to approximate the transition amplitude by building an estimator for the RHS, sampling over classical configurations, and summing up the estimator.

Now, the expression for U was not-so-pretty, but imagine continuous time variables and integrals when you see $\sum_k$ and ϵ above. In the limit as $N \to \infty$ (the number of completeness relations inserted), the quantum propagator is expressed in a continuous form with strange new "integrals".

$$U(q_i, q_f; T) = \left(\int \mathcal{D}q \int \mathcal{D}p \right) e^{i \int_0^T dt \, (\sum_j p^j \dot{q}^j - H(q,p))} \tag{17}$$

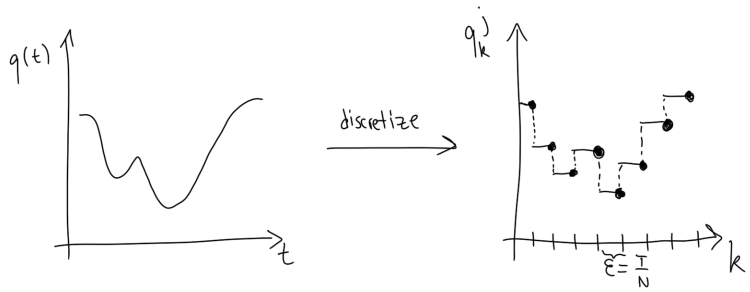
Do not think of these as literal integrals, as we do not have a proper measure space to integrate over. Think of them as algorithms for now, something totally new that will be applied to solve this expression above.

Lecture 2: Gaussian Path Integrals

Recall the propagator, or transition amplitude, for a nonrelativistic quantum system

$$U(q_i, q_f; T) = \left(\prod_j \int \mathcal{D}q^j(t) \int \mathcal{D}p^j(t) \right) e^{i \int_0^T dt \, \mathcal{L}(q^j, \dot{q}^j)}. \tag{18}$$

To work with this, we often discretize $q(t) \to q_k^j$



$$U(q_i, q_f; T) = \left(\prod_{j,k} \int dq_k^j \int \frac{dp_k^j}{2\pi} \right) e^{i \sum_k (\sum_j p_k^j (q_{k+1}^j - q_k^j) - \epsilon H)} \tag{19}$$

Evaluate these very many integrals to get an answer dependent on $\epsilon = \frac{T}{N}$, since we discretized, take the limit as $\epsilon \to 0$ and deal with any encountered infinities.

Key Example

Consider the classical Hamiltonian

$$H = \frac{p^2}{2m} + V(q). \tag{20}$$

Calculate the transition amplitude (**Exercise**)

$$U(q_i, q_f; T) = \left(\prod_{j,k} \int dq_k^j \int \frac{dp_k^j}{2\pi} \right) e^{i \sum_k (\sum_j p_k^j (q_{k+1}^j - q_k^j) - \epsilon H)} \tag{21}$$

$$= \left(\prod_k \int dq_k \int \frac{dp_k}{2\pi} \right) e^{i \sum_k (p_k (q_{k+1} - q_k) - \epsilon (\frac{p_k^2}{2m} + V(q)))} \tag{22}$$

$$= \left(\prod_k \int dq_k \right) \sqrt{\frac{-im}{2\pi\epsilon}} e^{i \sum_k \frac{m}{2\epsilon} (q_{k+1} - q_k)^2 - \epsilon V\left(\frac{q_{k+1} + q_k}{2}\right)}. \tag{23}$$

We may also write this in the following notation, using the fact that the argument of the exponential is the discretized version of the action, now without the p-integral

$$\lim_{\epsilon \to 0} U(q_i, q_f; T) = \int \mathcal{D}q(t) e^{\mathcal{S}[q(t)]} \tag{24}$$

Where the action is

$$\mathcal{S}[q(t)] = \int_0^T dt \ \left(\frac{m}{2} \sum_j (\dot{q}^j)^2 - V(q) \right). \tag{25}$$

Note that if our system is a harmonic oscillator $V(q) = \frac{1}{2} m\omega^2 q^2$, we can do the full integral.

Path Integrals for Scalar Fields

Recall the classical scalar field with Lagrangian density and Hamiltonian

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - V(\phi) \tag{26}$$

$$H = \int d^3x \ \left(\frac{1}{2}\pi^2(x) + \frac{1}{2}(\nabla \phi(x))^2 + V(\phi(x)) \right). \tag{27}$$

The path integral prescription for quantum scalar fields gives the transition amplitude, by blind application of the above, we conjecture that

$$\langle \phi_b | \, e^{-i\hat{H}T} \, | \phi_a \rangle = \left(\int \mathcal{D}\phi \int \mathcal{D}\pi \right) e^{i \int_0^T d^4x \, (\pi \dot{\phi} - H(\phi))} \tag{28}$$

Where the boundary terms are $\phi(t = 0, x) = \phi_a(\mathbf{x})$ and $\phi(t = T, x) = \phi_b(\mathbf{x})$.

As explained above, to make sense of this quantity, we must discretize, evaluate, and take the continuum limit as $\epsilon \to \infty$. When we discretize, note that we only discretize space, as discretizing time in this way will cause trouble with the conjugate momenta.

The field operators are discretized over a "grid" of points x_j each of width ϵ, such that

$$\phi(t, x) \ \ \to \ \ \phi(t, x_j) \equiv q^j(t). \tag{29}$$

Then discretize the integral by turning it into a sum over the grid

$$\int d^3x \ \ \to \ \ \epsilon^3 \Sigma_{j \in \mathbb{Z}^3}. \tag{30}$$

Next the derivative can be discretized via a finite difference. Note that there are more computationally efficient symmetric differences that can be used to discretize the derivative, but the finite difference works well for demonstration

$$\nabla_\mu \phi(x) \;\; \rightarrow \;\; \frac{(\phi(x_j + \epsilon_\mu) - \phi(x_j))}{|\epsilon_\mu|} \tag{31}$$

Where μ denotes the four directions in which to calculate the derivative

$$\epsilon_\mu = \epsilon\{\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix}\} \tag{32}$$

.
Lastly, the potential just becomes evaluated at each x_j

$$V(\phi(x)) \;\; \rightarrow \;\; V(\phi(x_j)). \tag{33}$$

Then the Lagrangian is discretized to a sum over a bunch of terms, but the only relevant term to the construction of the Hamiltonian is the time derivative of the field operator $\dot{\phi}$ (**Exercise**)

$$L = \int d^3x \, \mathcal{L} \rightarrow \epsilon^3 \sum_j \frac{1}{2}(\dot{\phi}_j)^2 \tag{34}$$

And the discretized conjugate momentum becomes

$$\pi^j = \frac{\partial L}{\partial \dot{q}^j} = \frac{\partial L}{\partial \dot{\phi}^j} \;\; \rightarrow \;\; \epsilon^3 \dot{q}^j. \tag{35}$$

Finally, we have the discretized Hamiltonian, where we display the ϵ terms to show that if we did not add the ϵ^3 term to the discretized Lagrangian, we would be stuck with an extra ϵ^{-3} on the discretized Hamiltonian

$$H = \epsilon^3 \sum_j \epsilon^{-3} \pi_j^2 + \frac{1}{2}\left(\frac{q_{j+\epsilon^\mu} - q_j}{\epsilon}\right)^2 + V(q). \tag{36}$$

In summary, the discretization of the scalar field gives us a nonrelativistic lattice system such that the discretized Hamiltonian is the sum of a kinetic energy term and a potential energy term. The second step is to evaluate the (nonrelativistic) path integral, and the third step is to take the continuum limit as $\epsilon \rightarrow 0$, which will later be re-branded as renormalization.

The most important case of the scalar field is the quadratic potential, which corresponds to the Klein-Gordon field (e.g., discretizing Klein-Gordon theory yields the quadratic potential below), is

$$V(q) = \frac{1}{2} q^T \mathbf{A} q. \tag{37}$$

Gaussian Integrals

Consider the following integral

$$I = \int_{-\infty}^{\infty} dx\ e^{-x^2} = \sqrt{\pi}. \tag{38}$$

Proof:

$$I = \int_{-\infty}^{\infty} dx\ e^{-x^2} \tag{39}$$

$$I^2 = \left(\int_{-\infty}^{\infty} dx\ e^{-x^2}\right)\left(\int_{-\infty}^{\infty} dy\ e^{-y^2}\right) \tag{40}$$

$$I^2 = \int_0^{\infty} r dr \int_0^{2\pi} d\theta\, e^{-r^2} \tag{41}$$

$$I^2 = 2\pi \int_0^{\infty} \frac{d}{dr}\left(-\frac{1}{2}e^{-r^2}\right) dr = \pi \tag{42}$$

.
This is actually a special case of the more general forms of the Gaussian integral

$$\int_{-\infty}^{\infty} dx\ e^{-\frac{1}{2}ax^2+bx} = \sqrt{\frac{2\pi}{a}} e^{\frac{b^2}{2a}} \tag{43}$$

$$\int_{-\infty}^{\infty} dx\ e^{iax^2+ibx} = \sqrt{\frac{2\pi i}{a}} e^{\frac{-ib^2}{2a}} \tag{44}$$

$$\tag{45}$$

.
We will later need the moments generated by the Gaussian integral

$$\langle x^n \rangle = \frac{\int_{-\infty}^{\infty} dx\ x^n e^{-\frac{1}{2}ax^2}}{\int_{-\infty}^{\infty} dx\ e^{-\frac{1}{2}ax^2}}. \tag{46}$$

Note that if n is odd, then the moment is zero and we can write the exponent of x as $2m$, where $m \in \mathbb{Z}$, and we have the relation (**Exercise**)

$$\langle x^{2m} \rangle = \frac{(2m-1)!!}{a^m}. \tag{47}$$

Note that the double factorial $(2m-1)!!$ represents the number of ways to join $2m$ points in pairs. – "All science should in linear algebra or combinatorics." –

Another closed form of this integral is in terms of derivatives is

$$\langle x^{2m} \rangle = \left(\frac{d}{db} \right)^{2m} \left(\frac{\int_{-\infty}^{\infty} dx \ e^{-\frac{1}{2}ax^2 + bx}}{\int_{-\infty}^{\infty} dx \ e^{-\frac{1}{2}ax^2}} \right) \Big|_{b=0} \tag{48}$$

$$\langle x^{2m} \rangle = \left(\frac{d}{db} \right)^{2m} e^{\frac{b^2}{2a}} \Big|_{b=0}. \tag{49}$$

To evaluate the maultivariable Gaussian integrals, where $x \in \mathbb{R}^n$, consider

$$I(\mathbf{A}, B) = \int_{-\infty}^{\infty} dx_1 \cdots \int_{-\infty}^{\infty} dx_n \ e^{-x^T \mathbf{A} x + B^T x} \tag{50}$$

Where $\mathbf{A}$ is an $n \times n$ symmetric real matrix and B is an $n \times 1$ real vector. Since $\mathbf{A}$ is real, symmetric, it contains orthogonal $\mathbf{O}$ and diagonal matrices $\mathbf{D}$, such that $\mathbf{O}^T\mathbf{O} = \mathbb{I}$ and $\mathbf{D}$ is diagonalized with the eignevalues of $\mathbf{A}$.

$$\mathbf{O}^T \mathbf{D} \mathbf{O} = \mathbf{A} \tag{51}$$

Assume that $B = 0$ and define $y = \mathbf{O}x$. Then

$$I(\mathbf{A}, B = 0) = \int_{-\infty}^{\infty} dy_1 \cdots \int_{-\infty}^{\infty} dy_n \ e^{-y^T \mathbf{D} y} \tag{52}$$

$$= \prod_{j=1}^{n} \int_{-\infty}^{\infty} dy_j \ e^{-y_j^2 \lambda_j} \tag{53}$$

$$= \prod_{j=1}^{n} \sqrt{\frac{\pi}{\lambda_j}} \tag{54}$$

$$I(\mathbf{A}, B = 0) = \sqrt{\frac{\pi^n}{\det(\mathbf{A})}}. \tag{55}$$

The $B \neq 0$ case (**Exercise**) results in the following

$$I(\mathbf{A}, B) = \sqrt{\frac{\pi^n}{\det(\mathbf{A})}} e^{B^T \mathbf{A}^{-1} B}. \tag{56}$$

Lecture 3: Correlation Functions and Path Integrals

Recall the generating function for a single-variable Gaussian probability distribution $e^{\frac{1}{2}ax^2}$ and the moment-generating integral

$$I = \int_{-\infty}^{\infty} dx\ x^{2n} e^{\frac{1}{2}ax^2} = \frac{(2n-1)!!}{a^n}. \tag{57}$$

We also derived the identity with the generating function for the multivariable Gaussian probability distribution

$$\int_{-\infty}^{\infty} dx_1 \cdots \int_{-\infty}^{\infty} dx_n\ e^{-\frac{1}{2}x^T \mathbf{A} x + J^T x} = \sqrt{\frac{\pi^n}{\det(\mathbf{A})}} e^{J^T \mathbf{A}^{-1} J}. \tag{58}$$

The 2-point correlation function for the n-variable Gaussian is (**Exercise**), for $j \neq k$

$$\langle x_j x_k \rangle \equiv \frac{\int_{-\infty}^{\infty} dx_1 \cdots \int_{-\infty}^{\infty} dx_n\ x_j x_k e^{-\frac{1}{2}x^T \mathbf{A} x}}{\int_{-\infty}^{\infty} dx_1 \cdots \int_{-\infty}^{\infty} dx_n\ e^{-\frac{1}{2}x^T \mathbf{A} x}} \equiv [\mathbf{A}^{-1}]_{jk}. \tag{59}$$

Note that this is also equal to the second derivative with respect to the vector J, evaluated at $J = 0$

$$\langle x_j x_k \rangle \equiv \frac{\frac{\partial^2}{\partial J_j \partial J_k} \int_{-\infty}^{\infty} dx_1 \cdots \int_{-\infty}^{\infty} dx_n\ e^{-\frac{1}{2}x^T \mathbf{A} x + J^T x}}{\int_{-\infty}^{\infty} dx_1 \cdots \int_{-\infty}^{\infty} dx_n\ e^{-\frac{1}{2}x^T \mathbf{A} x + J^T x}}\Big|_{J=0}. \tag{60}$$

Higher Order Moments: l-point Correlation Functions

The l-point correlation function has similar form

$$\langle x_{j_1} \dots x_{j_l} \rangle \equiv \frac{\int_{-\infty}^{\infty} dx_{j_1} \cdots \int_{-\infty}^{\infty} dx_{j_l}\ x_{j_1} \dots x_{j_l}\ e^{-\frac{1}{2}x^T \mathbf{A} x}}{\int_{-\infty}^{\infty} dx_{j_1} \cdots \int_{-\infty}^{\infty} dx_{j_l}\ e^{-\frac{1}{2}x^T \mathbf{A} x}}. \tag{61}$$

By Wick's theorem (proof by induction), for even l, the l-point correlation function is proportional to the sum of the products over the permutation group on l symbols, the "Wick sum". We write "proportional to" for reasons of symmetry and soon eliminating redundant terms.

$$\langle x_{j_1} \dots x_{j_l} \rangle \propto \sum_{\pi \in S_l} [\mathbf{A}^{-1}]_{j_{\pi^{-1}(1)} j_{\pi^{-1}(2)}} \dots [\mathbf{A}^{-1}]_{j_{\pi^{-1}(l-1)} j_{\pi^{-1}(l)}}. \tag{62}$$

Example: 4-point Correlation

To calculate the 4-point correlation function, we begin by considering the $4! = 24$ total permutations on 4 symbols. Since $\mathbf{A}^{-1}$ is symmetric

$$[\mathbf{A}^{-1}]_{jk} = [\mathbf{A}^{-1}]_{kj} \tag{63}$$

And there are only $\frac{24}{2!2!2!} = 3$ unique terms (products of two matrix elements), which are (**Exercise**)

$$\begin{aligned}\langle x_{j_1} x_{j_2} x_{j_3} x_{j_4} \rangle &= [\mathbf{A}^{-1}]_{j_1 j_2} [\mathbf{A}^{-1}]_{j_3 j_4} \\ &+ [\mathbf{A}^{-1}]_{j_1 j_3} [\mathbf{A}^{-1}]_{j_2 j_4} \\ &+ [\mathbf{A}^{-1}]_{j_1 j_4} [\mathbf{A}^{-1}]_{j_2 j_3} \\ &= \langle x_{j_1} x_{j_2} \rangle \langle x_{j_3} x_{j_4} \rangle + \langle x_{j_1} x_{j_3} \rangle \langle x_{j_2} x_{j_4} \rangle + \langle x_{j_1} x_{j_4} \rangle \langle x_{j_2} x_{j_3} \rangle\end{aligned}$$

With the approporiate choice of $\mathbf{A}$, in the context of path integrals and perturbative field theory, these products of correlations functions are exactly correspondent to Feynman propagator, and, in turn, the Feynman diagrams, just as we studied in *Lecture 9* of the last lecture series (Quantum Field Theory) for the 4-particle Wick contraction.

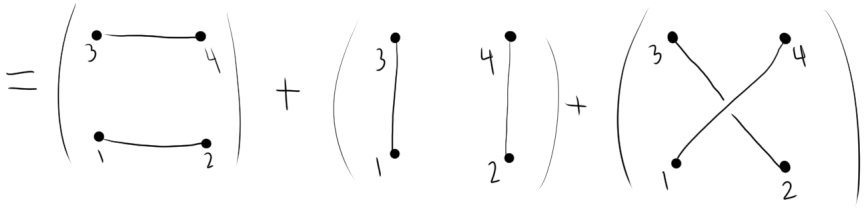


Figure 1: Feynman diagram correspondence of the 4-point correlation function

Keeping only unique terms, the proprotionality relation becomes an equivalence

$$\langle x_{j_1} \dots x_{j_l} \rangle = \sum_{\text{unique } \pi^{-1}} [\mathbf{A}^{-1}]_{j_{\pi^{-1}(1)} j_{\pi^{-1}(2)}} \dots [\mathbf{A}^{-1}]_{j_{\pi^{-1}(l-1)} j_{\pi^{-1}(l)}} \tag{64}$$

(**Exercise**) Calculate the 6-point correlation function with $\frac{6!}{2!2!2!} = 90$ unique terms

$$\langle x_{j_1} x_{j_2} x_{j_3} x_{j_4} x_{j_5} x_{j_6} \rangle = [\mathbf{A}^{-1}]_{j_1 j_2} [\mathbf{A}^{-1}]_{j_3 j_4} [\mathbf{A}^{-1}]_{j_5 j_6} + \dots \tag{65}$$

In short summary,

- We can calculate *all* moments for the Gaussian probability distribution.
- We have a diagrammatic calculus to calculate the l-point correlation functions, which end up being exactly the Feynman propagators/diagrams, with appropriate choice of $\mathbf{A}$, and is the direct connection of quantum field theory and Gaussian integrals.

The Matrix A for Path Integrals

Let the potential V be quadratic in the canonical position coordinate per particle q_k (e.g., a one-dimensional chain of oscillators), such that the transition amplitude, which will be discretized, evaluated, and limited $\epsilon \to 0$, from some state q_a to another q_b is

$$U(q_a, q_b; T) = \left(\prod_k \int \frac{dq_k}{c(\epsilon)} \right) e^{\frac{1}{2} i q^T \mathbf{A} q} \tag{66}$$

Where we know the quadratic form contains a kinetic energy term plus a potential energy term

$$q^T \mathbf{A} q = \sum_k \left(m \frac{(q_{k+1} - q_k)^2}{\epsilon} - \epsilon V(\frac{q_{k+1} + q_k}{2}) \right). \tag{67}$$

This results is $\mathbf{A}$ as a *tridiagonal* matrix for the kinetic energy term and a potential energy term which is a matrix with elements quadratic in q

$$\mathbf{A} = \begin{bmatrix} \frac{2m}{\epsilon} & -\frac{m}{\epsilon} & 0 & \cdots & & \\ -\frac{m}{\epsilon} & \frac{2m}{\epsilon} & -\frac{m}{\epsilon} & 0 & \cdots & \\ 0 & -\frac{m}{\epsilon} & \frac{2m}{\epsilon} & -\frac{m}{\epsilon} & 0 & \cdots \\ \vdots & 0 & \ddots & \ddots & \ddots & \end{bmatrix} + [V(q^2)] \tag{68}$$

The transition amplitude is then calculated similarly to last lecture as

$$U(q_a, q_b; T) = \frac{\infty \, \text{const.}}{\sqrt{\det(\mathbf{A})}} \tag{69}$$

The infinite constant will not be a problem since the l-point correlation is normalized, and the same exact infinite constant will appear in the denominator and cancel the constant. So, in terms of q_k, the l-point correlation reads

$$\langle q_{j_1} \dots q_{j_l} \rangle \equiv \frac{\prod_k \int \frac{dq_k}{c(\epsilon)} q_{j_1} \dots q_{j_l} \; e^{-\frac{1}{2} i q^T \mathbf{A} x}}{\prod_k \int \frac{dq_k}{c(\epsilon)} \; e^{-\frac{1}{2} i q^T \mathbf{A} x}}. \tag{70}$$

Note that with periodic boundary conditions, the elements of follow a modulo relation $\mathbf{A}_{jk} = f((j-k) \bmod n)$, where n is the number of sites/oscillators in the chain.

Assuming that $\mathbf{A}$ is invertible, there exists a unitary matrix Q, such that $Q^T Q = \mathbb{I}$ and $Q^T \mathbf{A} Q = D$, with diagonal matrix D with the eigenvalues of $\mathbf{A}$ along the diagonal. Then the determinant of $\mathbf{A}$ is easy to calculate, since

$$\det(\mathbf{A}) = \prod_{j=1}^{n} \lambda_j(\mathbf{A}). \tag{71}$$

Correlations Functions and Quantum Observables

Consider the transition amplitude over 2-point spatial correlations

$$U(q_a, q_b; T) \propto \int \mathcal{D}\phi(x)\ \phi(x_1)\phi(x_2) e^{i\int_{-T}^{T} d^4x\, \mathcal{L}(\phi)} \tag{72}$$

With an expression like this, always discretize by sending $\phi(x_j) \to q_j$, evaluate the integral, and enter the contiuum limit with the boundary conditions

$$\phi(-T, x) = \phi_a(x) \tag{73}$$
$$\phi(T, x) = \phi_b(x). \tag{74}$$

Apply the following condition, exploiting the boundary conditions, to factor the full field "integral" over the individual fields and the boundary of the field

$$\int \mathcal{D}\phi(x) = \int \mathcal{D}\phi_1(x) \int \mathcal{D}\phi_2(x) \int_{\partial\phi} \mathcal{D}\phi(x) \tag{75}$$

Where the boundary $\partial\phi = \partial\phi_1 + \partial\phi_2$ is defined by

$$\phi_1(x) = \phi(x_1^0, x_1) \tag{76}$$
$$\phi_2(x) = \phi(x_2^0, x_2) \tag{77}$$

So, after performing the boundary integral, we introduce quantum stuff to the expression from the classical 2-point function above, for $x_2^0 > x_1^0$

$$U(q_a, q_b; T) \propto \int \mathcal{D}\phi_1(x) \int \mathcal{D}\phi_2(x)\ \phi(x_1)\phi(x_2) \left\langle \phi_b \right| e^{-i\hat{H}(T - x_2^0)} \left| \phi_2 \right\rangle \\ \times \left\langle \phi_2 \right| e^{-i\hat{H}(x_2^0 - x_1^0)} \left| \phi_1 \right\rangle \left\langle \phi_1 \right| e^{-i\hat{H}(x_1^0 + T)} \left| \phi_a \right\rangle$$

Now, apply the Schroedinger-picture field operator to write the classical field operators in terms of quantum field operators. The formula is

$$\hat{\phi}_S(x) \left| \phi_1 \right\rangle = \phi(x_1) \left| \phi_1 \right\rangle \tag{78}$$

So, the purely quantum expression for the 2-point function is

$$U(q_a, q_b; T) \propto \int \mathcal{D}\phi_1(x) \int \mathcal{D}\phi_2(x)\ \left\langle \phi_b \right| e^{-i\hat{H}(T - x_2^0)} \hat{\phi}_S(x) \left| \phi_2 \right\rangle \\ \times \left\langle \phi_2 \right| e^{-i\hat{H}(x_2^0 - x_1^0)} \hat{\phi}_S(x) \left| \phi_1 \right\rangle \left\langle \phi_1 \right| e^{-i\hat{H}(x_1^0 + T)} \left| \phi_a \right\rangle$$

This is called the time-ordered expectation value of the field operators in the Heisenberg picture. The equation holds for $x_2^0 < x_1^0$ as well, and we can write it as

$$U(q_a, q_b; T) \propto \langle \phi_b | \, e^{-i\hat{H}T} \mathcal{T}[\hat{\phi}_H(x_1)\hat{\phi}_H(x_2)] e^{-i\hat{H}T} \, | \phi_a \rangle \tag{79}$$

Now, enter the limit as $T \to \infty$, where we bring the interacting vacuum state and the normalization for the full transition amplitude (**Exercise**), and introduce the *most important formula* for this course

$$\langle \Omega | \, \mathcal{T}[\hat{\phi}_H(x_1)\hat{\phi}_H(x_2)] \, | \Omega \rangle = \lim_{T \to \infty(1-i\epsilon)} \frac{\int \mathcal{D}\phi \; \phi(x_1)\phi(x_2) e^{i \int_{-T}^{T} d^4x \, \mathcal{L}(\phi)}}{\int \mathcal{D}\phi \; e^{i \int_{-T}^{T} d^4x \, \mathcal{L}(\phi)}}. \tag{80}$$

So, the LHS is built of purely quantum observables equal to the classical expression of path integrals!

This expression will end up to be the propagator, which is also the inverse of the Klein-Gordon operator, which is what we call **A** in the scalar quantum field theory.

The solution to this is well-known for the case when $\mathcal{L}$ is quadratic in the field operators, and one can easily discretize, evaluate the Gaussian integral, and take the limit as $\epsilon \to 0$.

(Exercise) Calculate the l-point formula for the time-ordered expectation value of the field operators in the Heisenberg picture.

Lecture 4: Functional Quantization of the Scalar Field

The path integral formalism for quantization of fields is an incredibly efficient tool, but one must learn when, and when not, ot use it. Through the lectures and many examples, we'll develop an intuition for when to trust quantization via path integrals.

Recall the action of the scalar field S with classical field operators ϕ

$$S_0 = \int d^4x \ \mathcal{L}_0 = \int d^4x \ \left(\frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2}m^2\phi^2\right). \tag{81}$$

We will (1) discretize, tantamount to imposing a cutoff Λ, (2) evaluate the integrals, and (3) enter the continuum limit where $\epsilon \to 0$. Start the discretization by putting the field on a lattice (a Lorentz manifold) with spacing ϵ and then compactify the space onto a torus for periodic boundary conditions.

Mathematically, we are transforming from a four-dimensional Minkowski space $\mathcal{M}^4$ to a four-domensional torus $(\mathbb{Z}/N\mathbb{Z})^4$, where $N = \frac{L}{\epsilon}$ is the number of sites, and L is the total size of grid.
Continue discretization with the field operators

$$\phi(x) \to \phi(x_j) \equiv q_j, \tag{82}$$

$$x_j = \epsilon j \in \frac{L}{N}(\mathbb{Z}/N\mathbb{Z})^4. \tag{83}$$

And the partial derivatives are replaced for the forward difference, which is not the best method, but it's "good enough for government work"

$$\partial_\mu \phi(x) \to \frac{\phi(x_j + \epsilon e^\mu) - \phi(x_j)}{\epsilon}. \tag{84}$$

And the space-time integral becomes a sum over the sites on the torus

$$\int d^4x \ \to \epsilon^4 \sum_{j \in (\mathbb{Z}/N\mathbb{Z})^4} \tag{85}$$

Now following the path integral quanitzation recipe, consider the transition amplitude in terms of the discretized action

$$\langle \phi_f | \, U(q_i, q_f; T) \, |\phi_i\rangle = \int \mathcal{D}\phi \ e^{iS_0} \tag{86}$$

Where we follwo the "algorithm" of the integral-differential operator and discretize it to a product, over the torus sites, of integrals (N^4 total integrals) over the field operators, the canonical position variables

$$\int \mathcal{D}\phi \to \prod_j \int d\phi(x_j) \equiv \prod_j \int dq_j. \tag{87}$$

Discretization in Momentum Space

Thus far we have worked entirely in real (position) space. Let's Fourier transform over into momentum space to continue discretization. The Fourier transform is a unitary transformation with Jacobian equal to 1 (**Exercise**). First, the field operators transform as

$$\phi(x_j) = \frac{1}{V} \sum_n e^{-ik_n \cdot x_j} \phi(k_n) \tag{88}$$

Where $V = L^4$ is the volume of the 4D torus. Notationally, the k argument to the field operator in momentum space $\phi(k)$ will denote the Fourier transform of the field operator in real space $\phi(x)$. The wavenumber k_n is discretized over the torus as

$$k_n = \frac{2\pi n^\mu}{L},\ n^\mu \in \mathbb{Z}/N\mathbb{Z} \text{ and } |k^\mu| < \frac{\pi}{\epsilon} \tag{89}$$

Note that the Fourier space field operator is complex, such that $\phi(-k) = \phi^*(k)$, and we therefore have two independent variables per field operator in momentum space: the real part $\Re\phi(k_n)$ and the imaginary part $\Im\phi(k_n)$ for positive time-component $k_n^0 > 0$.
So, in momentum space, the discretized integral-differential operator is (**Exercise**)

$$\int \mathcal{D}\phi = \prod_{n:k_n^0>0} \int d\,\Re\phi(k_n) \int d\,\Im\phi(k_n). \tag{90}$$

And the discretized action for the scalar field in momentum space is (**Exercise**)

$$S_0 = -\frac{1}{V} \sum_{k_n^0>0} (m^2 - k_n^2)((\Re\phi_n)^2 + (\Im\phi_n)^2) \tag{91}$$

Where $\phi_n \equiv \phi(k_n)$, and the following relation for the Kronecker delta is used to obtain this expression

$$\delta_{k,0} = \frac{1}{n} \sum_{j=0}^{n-1} e^{\frac{2\pi i j k}{n}}. \tag{92}$$

Our expression for the path integral for the Klein-Gordon field discretized to a lattice (four-dimensional with periodic boundary conditions) is comprised of Gaussian integarls over a finite number of degrees of freedom

$$I_0 = \int \mathcal{D}\phi\, e^{iS_0} = \left(\prod_{k_n^0>0} \int d\,\Re\phi_n \int d\,\Im\phi_n \right) e^{-i\frac{1}{V}\sum_{k_n^0>0}(m^2-k_n^2)|\phi_n|^2}. \tag{93}$$

Now, onto evaluating this integral, it's just a bunch of Gaussian integrals, and we know how to solve those. We get the following, and unrestrict k_n to get the second line (**Exercise**)

$$I_0 = \prod_{k_n^0>0} \sqrt{\frac{-i\pi V}{m^2 - k_n^2}} \cdot \sqrt{\frac{-i\pi V}{m^2 - k_n^2}} \tag{94}$$

$$I_0 = \prod_{k_n} \sqrt{\frac{-i\pi V}{m^2 - k_n^2}} \tag{95}$$

Note that k_n is bounded, but we have an infinity when $V \to \infty$ (continuum limit), but this integral does not yet have full operational meaning and is proportional to the transition amplitude $I_0 \propto \langle \phi_f | \, U(q_i, q_f; T) \, | \phi_i \rangle$, and the infinities will cancel and drop out in the full expression.

Heuristic Argument for I_0

As the "surface area of knowledge" we need to remember the path integral formalism is small, there is a heuristic way to obtain this result without formal discretization, etc., using the aforementioned intuition.

Recall the Gaussian integral whose argument is quadratic in its independent variable

$$\int dx \; e^{-x^T \mathbf{A} x} = \sqrt{\frac{\pi^n}{\det(\mathbf{A})}} \propto (\det(\mathbf{A}))^{-\frac{1}{2}} \tag{96}$$

For the Klein-Gordon field, consider the path integral with the Klein-Gordon operator and field operators substituted

$$\int \mathcal{D}\phi \; e^{iS} \; \sim \; \int \mathcal{D}\phi \; e^{\frac{1}{2}\int d^4x \; \phi(x)(-\partial^2 - m^2)\phi(x)} \tag{97}$$

So, we are boldly extrapolating to say that $\mathbf{A}$ is like the Klein-Gordon operator and the x is like the field operator

$$\int d^4x \; \phi(x)(-\partial^2 - m^2)\phi(x) \sim x^T \mathbf{A} x. \tag{98}$$

Furthermore, we say that the path integral is proportional to the determinant of the Klein-Gordon operator

$$\int \mathcal{D}\phi \; e^{iS} \; \propto \; (\det(-\partial^2 - m^2))^{-\frac{1}{2}}. \tag{99}$$

Operationally Well-Defined Quantities

As mentioned, I_0 cancels for operationally well-defined quanities, such as the 2-point correlation function, a time-ordered expectation value of products of the field operators. For example, using the path integral formalism

$$\langle\Omega|\,\mathcal{T}[\phi(x_1)\phi(x_2)]\,|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\int \mathcal{D}\phi\ \phi(x_1)\phi(x_2)e^{iS}}{\int \mathcal{D}\phi\ e^{iS}} \tag{100}$$

To check our bold extrapolations, calculate the discretized field operator product

$$\phi(x_1)\phi(x_2) = \frac{1}{V^2}\sum_m e^{-ik_m\cdot x_1}\phi_m \sum_l e^{-ik_l\cdot x_2}\phi_l \tag{101}$$

So, the discretized RHS numerator of the time-ordered expectation value above is just a bunch of independent Gaussian integrals, quadratic in its independent variables (**Exercise**)

$$\begin{aligned}
\text{numerator} &= \frac{1}{V^2}\sum_{l,m} e^{-i(k_m\cdot x_1 + k_l\cdot x_2)} \left(\prod_{k_n^0>0}\int d\,\Re\phi_n \int d\,\Im\phi_n\right) \\
&\quad\times (\Re\phi_m + i\Im\phi_m)(\Re\phi_l + i\Im\phi_l) e^{-i\frac{1}{V}\sum_{k_n^0>0}(m^2-k_n^2)((\Re\phi_n)^2+(\Im\phi_n)^2)} \\
&= \frac{1}{V^2}\sum_m e^{-ik_m\cdot(x_1-x_2)} \left(\prod_{k_n^0>0}\frac{-i\pi V}{m^2-k_n^2}\right)\frac{-iV}{m^2-k_n^2-i\epsilon} \\
&= \frac{1}{V^2}\sum_m e^{-ik_m\cdot(x_1-x_2)}\cdot I_0\cdot\frac{-iV}{m^2-k_n^2-i\epsilon}
\end{aligned}$$

Where we drastically cut down the number of integrals to evaluate, since any integrals involving products like $\Re\phi_m\cdot\Im\phi_l$ or $\Im\phi_m\cdot\Re\phi_l$ form odd integrands and evaluate to zero. The integeral will also be zero for terms where $m\neq l$ and for terms where $k_m = k_l$. Integrals where $k_m = -k_l$ will *not* be zero.

Bringing this together, the RHS of the time-ordered expectation value has boiled down to

$$\langle\Omega|\,\mathcal{T}[\phi(x_1)\phi(x_2)]\,|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\int \mathcal{D}\phi\ \phi(x_1)\phi(x_2)e^{iS}}{\int \mathcal{D}\phi\ e^{iS}} \tag{102}$$

$$= \lim_{V\to\infty} -i\frac{1}{V}\sum_n \frac{e^{-ik_n\cdot(x_1-x_2)}}{m^2-k_n^2-i\epsilon} \tag{103}$$

$$= \int \frac{d^4k}{(2\pi)^4}\ \frac{ie^{-ik\cdot(x_1-x_2)}}{-m^2+k^2+i\epsilon} \tag{104}$$

$$\langle\Omega|\,\mathcal{T}[\phi(x_1)\phi(x_2)]\,|\Omega\rangle = D(x_1-x_2) \tag{105}$$

So, the path integral formalism gives us exactly the propagator we wish to see. Note that is we were to just boldy extrapolate, without discretization, etc., we would get the same result!

For example,

$$\frac{\int \mathcal{D}\phi\ \phi(x_1)\phi(x_2)e^{iS}}{\int \mathcal{D}\phi\ e^{iS}} = \frac{(\partial^2 - m^2)^{-\frac{1}{2}} D(x_1 - x_2)}{(\partial^2 - m^2)^{-\frac{1}{2}}} \tag{106}$$

Since if $\mathbf{A} \sim (-\partial^2 - m^2)$
Then $[\mathbf{A}^{-1}]_{jk} \sim \frac{1}{(-\partial^2 - m^2)_{x_1 x_2}} = D(x_1 - x_2)$
And $\delta^{(4)}(x - y) = (-\partial^2 - m^2)D(x - y)$.

Example: 4-point Correlation Function

Note that all 3-point correlations are zero, since they all have odd integrands. The 4-point correlation function starts off as

$$\langle\Omega|\,\mathcal{T}[\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)]\,|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\int \mathcal{D}\phi\ \phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)e^{iS}}{\int \mathcal{D}\phi\ e^{iS}}. \tag{107}$$

The numerator contains the quantities of $(\Re\phi_m + i\Im\phi_m)\ldots(\Re\phi_l + i\Im\phi_l)$, and most terms will vanish as before, leaving us with terms where $k_l = -k_m$ and $k_q = -k_p$, and we end up with, after applying Wick's theorem and sending $V \to \infty$ (**Exercise**), something like

$$\sum_{k_l,k_q} e^{-i\ldots} \int \ldots \phi_{k_l}\phi_{-k_l}\phi_{k_q}\phi_{-k_q} e^{\ldots}$$
$$= D_F(x_1 - x_2)D_F(x_3 - x_4) + D_F(x_1 - x_3)D_F(x_2 - x_4) + D_F(x_1 - x_4)D_F(x_2 - x_3)$$

Interacting QFT via Path Integrals

Consider the action with a free part and an interacting part, namely the phi-fourth interaction,

$$S = S_0 + S_{int} = S_0 + \frac{i\lambda}{4!}\int d^4x\ \phi^4(x). \tag{108}$$

Then the time-ordered expectation value for 2-point correlations can be Taylor expanded, since λ is small,

$$\langle\Omega|\,\mathcal{T}[\phi(x_1)\phi(x_2)]\,|\Omega\rangle = lim_{T\to\infty(1-i\epsilon)}\frac{\int\mathcal{D}\phi\ \phi(x_1)\phi(x_2)e^{i(S_0+S_{int})}}{\int\mathcal{D}\phi\ e^{i(S_0+S_{int})}} \tag{109}$$

$$= lim_{T\to\infty(1-i\epsilon)}\frac{\int\mathcal{D}\phi\ \phi(x_1)\phi(x_2)e^{iS_0}(1+S_{int}+\frac{1}{2}S_{int}^2+\dots)}{\int\mathcal{D}\phi\ e^{iS_0}(1+S_{int}+\frac{1}{2}S_{int}^2+\dots)} \tag{110}$$

Where $S_{int} = \frac{i\lambda}{4!}\int d^4z\ \phi^4(z)$, and each term above is an integral of powers of time-ordered quantum field operators which end up as Feynman diagrams, for example, of the form

$$\frac{\lambda^m}{4!^m}\int d^4z_1\cdots\int d^4z_m\int\mathcal{D}\phi\ \phi(x_1)\phi(x_2)\phi^4(z_1)\dots\phi^4(z_m)e^{iS_0} \tag{111}$$

$$= \frac{\lambda^m}{4!^m}\int d^4z_1\cdots\int d^4z_m\,\langle\Omega|\,\mathcal{T}[\hat{\phi}(x_1)\hat{\phi}(x_2)\hat{\phi}^4(z_1)\dots\hat{\phi}^4(z_m)]\,|\Omega\rangle \tag{112}$$

$$= \text{Sum of Feynman diagrams} \tag{113}$$

Lecture 5: Functional Derivatives and Generating Functionals

Here we will finish the functional quantization of the scalar field.

Recall that we can compute time-ordered correlation functions for the quantum scalar field entirely in terms of classical quantities, which is equivalent to a sum over all diagrams,

$$\langle\Omega|\, T[\hat{\phi}(x_1)\dots\hat{\phi}(x_n)]\,|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\int \mathcal{D}\phi\ \phi(x_1)\dots\phi(x_n)e^{iS[\phi(x_1),\dots,\phi(x_n)]}}{\int \mathcal{D}\phi\ e^{iS[\phi(x_1),\dots,\phi(x_n)]}} \tag{114}$$

For example, the 2-point correlation function for the Klein-Gordon field is the Feynman propagator

$$\langle\Omega|\, T[\hat{\phi}(x_1)\hat{\phi}(x_2)]\,|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\int \mathcal{D}\phi\ \phi(x_1)\phi(x_2)e^{iS}}{\int \mathcal{D}\phi\ e^{iS}} = D_F(x_1-x_2) \tag{115}$$

More elegantly, and analogous to multivariate Gaussian integrals, we found the Feynman propagator $D_F(x-y)$, which is the inverse operator of the Klein-Gordon operator $-\partial^2 - m^2$, to be similiar to the inverse of a matrix $\mathbf{A}$, making the Klein-Gordon operator the matrix $\mathbf{A}$.

$$D_F(x_j - x_k) \sim [\mathbf{A}^{-1}]_{jk} = \frac{\int dx_1\dots dx_n\ x_j x_k e^{-\frac{1}{2}x^T\mathbf{A}x}}{\int dx_1\dots dx_n\ e^{-\frac{1}{2}x^T\mathbf{A}x}} \tag{116}$$

To compute these n-point correlation functions, or elements of this "inverse matrix", we used derivatives of the generating functional, which is what we generalize in this lecture. Recall the multivariate Gaussian generating functional

$$Z[J] = \int dx_1\dots dx_n e^{-\frac{1}{2}x^T\mathbf{A}x - J^Tx} = e^{\frac{1}{2}J^T\mathbf{A}^{-1}J}. \tag{117}$$

Functional Derivatives

The functional derivative is a tool from the calculus of variations that we now define by an example that is the continuum analog of the standard partial derivative

$$\frac{\delta}{\delta J(x)}J(y) = \delta^{(4)}(x-y). \tag{118}$$

There is a *one-to-one* mapping from the discrete representation to the continuous with correspondences

$$\begin{array}{ccc} x \in \mathbb{R} & \to & j \in \mathbb{Z} \\ J(x) \in C(\mathbb{R}) & \to & J_j \in L_2(\mathbb{Z}) \\ \frac{\delta}{\delta J(x)} F[J(y)] & \to & \frac{\partial}{\partial J_j} F[J_1, J_2, \dots] \end{array} \tag{119}$$

Where $C(\mathbb{R})$ is a continuous function space and $L_2(\mathbb{Z})$ is ...

Example 1

$$\begin{aligned} \frac{\delta}{\delta J(x)} e^{i\int d^4y\, J(y)\phi(y)} &= ie^{i\int d^4y\, J(y)\phi(y)} \frac{\delta}{\delta J(x)} \left(\int d^4y\, J(y)\phi(y) \right) \\ &= ie^{i\int d^4y\, J(y)\phi(y)} \int d^4y\, \frac{\delta J(y)}{\delta J(x)} \phi(y) \\ &= ie^{i\int d^4y\, J(y)\phi(y)} \int d^4y\, \delta^{(4)}(x-y)\phi(y) \\ \frac{\delta}{\delta J(x)} e^{i\int d^4y\, J(y)\phi(y)} &= i\phi(x) e^{i\int d^4y\, J(y)\phi(y)} \end{aligned}$$

Example 2: Derivatives of Delta functions

$$\begin{aligned} \frac{\delta}{\delta J(x)} \int d^4y\, (\partial_\mu J(y)) v^\mu(y) &= \frac{\delta}{\delta J(x)} \left(\text{boundary term} - \int d^4y\, J(y) \partial_\mu v^\mu(y) \right) \\ &= -\partial_\mu v^\mu(x) \end{aligned}$$

Note that the boundary term is almost always zero, except for topologically interesting theories.

The Generating Functional

Define the generating functional as

$$Z[J] = \lim_{T \to \infty(1-i\epsilon)} \int \mathcal{D}\phi\; e^{iS + iJ(x)\phi(x)}. \tag{120}$$

This expression is obviously useful, since correlation functions are directly related to derivatives of $Z[J]$

$$\langle \Omega |\, T[\hat{\phi}(x)\hat{\phi}(y)]\, | \Omega \rangle = \frac{-\frac{\delta}{\delta J(x)} \frac{\delta}{\delta J(y)} Z[J]\big|_{J=0}}{Z[J]\big|_{J=0}} \tag{121}$$

So, if you can compute the generating functional $Z[J]$, you have *all* of the n-point correlation functiosn via derivatives for your field theory.

In free field theories, such as the Klein-Gordon field, the action is quadratic in the field operators, and the argument of the exponential is $Z[J]$ is

$$i(S_0 + J(x)\phi(x) = i\int d^4x \ \left(\frac{1}{2}\phi(x)(-\partial^2 - m^2 + i\epsilon)\phi(x) + J(x)\phi(x)\right). \quad (122)$$

To homogenize quadraticity, complete the square by introducing the shift (with Jacobian = 1

$$\phi'(x) = \phi(x) - i\int d^4y \ D_F(x-y)J(y). \quad (123)$$

This is analogous to the positional shift $x' = x - \mathbf{A}^{-1}J$, and works becuase the Feynman propagator is the inverse of the Klein-Gordon operator, such that

$$(-\partial^2 - m^2)D_F(x-y) = i\delta^{(4)}(x-y) \quad (124)$$

With the variable change, the exponential argument becomes

$$\begin{aligned} i(S_0 + J(x)\phi(x) =& i\int d^4x \ \left(\frac{1}{2}\phi'(x)(-\partial^2 - m^2 + i\epsilon)\phi'(x)\right) \\ & - i\int d^4x \int d^4y \ \left(\frac{1}{2}J(x)(-iD_F(x-y))J(y)\right). \end{aligned}$$

So, the generating functional is then

$$Z[J] = Z_0 e^{-\frac{1}{2}\int d^4x d^4y \, (J(x)D_F(x-y)J(y))} \quad (125)$$

Where the free field contribution, independent of J, is

$$Z_0 = \int \mathcal{D}\phi' \ e^{i\int d^4x \, \left(\frac{1}{2}\phi'(x)(-\partial^2 - m^2 + i\epsilon)\phi'(x)\right)}. \quad (126)$$

Examples of Free Theory Correlations Functions

Example 1: The 2-point correlation function, with the Z_0 cancelled out,

$$\langle\Omega|\, T[\hat{\phi}(x)\hat{\phi}(y)]\,|\Omega\rangle = -\frac{\delta}{\delta J(x)}\frac{\delta}{\delta J(y)} e^{-\frac{1}{2}\int d^4x d^4y \, (J(x)D_F(x-y)J(y))}\Big|_{J=0}. \quad (127)$$

Example: The 4-point correlation function, with notation $\hat{\phi}_i = \hat{\phi}(x_i)$, $J_i = J(x_i)$, and $D_{xi} = D(x - x_i)$

$$\langle 0|\, T[\hat{\phi}_1\hat{\phi}_2\hat{\phi}_3\hat{\phi}_4]\,|0\rangle = \frac{\frac{\delta}{\delta J_1}\frac{\delta}{\delta J_2}\frac{\delta}{\delta J_3}\frac{\delta}{\delta J_4}Z[J]\big|_{J=0}}{Z[J=0]} \tag{128}$$

$$= \frac{\delta}{\delta J_1}\frac{\delta}{\delta J_2}\frac{\delta}{\delta J_3}\left(-\int d^4x'\ J_{x'}D_{x'4}e^{-\frac{1}{2}\int d^4x\int d^4y J_x D_{xy}J_y}\right)\big|_{J=0} \tag{129}$$

$$= \frac{\delta}{\delta J_1}\frac{\delta}{\delta J_2}\left(-D_{34}+\int d^4x'\int d^4y'\, J_{x'}D_{x'3}J_{y'}D_{y'4}\right)\times e^{\cdots}\big|_{J=0} \tag{130}$$

$$= \frac{\delta}{\delta J_1}\left(D_{34}\int d^4x'\, J_{x'}D_{x'2}+D_{24}\int d^4y'\, J_{y'}D_{y'3}+D_{23}\int d^4z'\, J_{z'}D_{z'4}+\mathcal{O}(J^2)\right)e^{\cdots}\big|_{J=0} \tag{131}$$

$$= D_{34}D_{12}+D_{24}D_{13}+D_{23}D_{14} \tag{132}$$

Interacting Fields

The time-ordered expectation value, which contains the generating functions by Taylor expansion, for the (classical) phi-fourth interacting theory is

$$\langle\Omega|\, T[\phi_1\dots\phi_n]\,|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)}\frac{\int\mathcal{D}\,\phi e^{i(S_0+S_{int})}\phi(x_1)\dots\phi(x_n)}{\int\mathcal{D}\phi\, e^{i(S_0+S_{int})}} \tag{133}$$

Where $S_{int}=-\frac{i\lambda}{4!}\int d^4x\,\phi^4(x)$.

Fermionic Fields

For the (classical) *fermionic field* $\hat{\psi}$, the 2-point correlation function, vacuum expectation value, is

$$\langle\Omega|\, T[\psi(x)\psi(y)]\,|\Omega\rangle = \frac{\int\mathcal{D}\,\psi e^{iS}\psi(x)\psi(y)}{\int\mathcal{D}\psi\, e^{iS}} \tag{134}$$

Rule number one for this expression (1) is to not think abou this operationally, and rule number two (2) is to think in analogy to complex numbers, which can provide a more clear representation and make things easier.

The Fermi fields obey the relations

$$\psi^2(x)=0 \tag{135}$$

$$\psi(x)\psi(y)=-\psi(y)\psi(x) \tag{136}$$

Vignette: Anticommuting Numbers (Grassman Numbers)

Let V be an n-dimensional vector space with basis $\theta_a \in V$, $a = 1, \ldots, n$. Thus, elements of the vector space $v \in V$ have the form

$$v = \sum_{a=1}^{n} v_a \theta_a. \tag{137}$$

To build a bigger vector space $\mathcal{G}(V)$ from V, we first endow V with the product operation denoted by concatenation (e.g., $\theta_a \cdot \theta_b \cdot \theta_c = \theta_a \theta_b \theta_c$).

This gives us an infinite dimensional vector space with span

$$\mathcal{S}^{\infty}(V) = \text{span}\{\theta_a, \theta_a\theta_b, \theta_a\theta_b\theta_c, \ldots\} \tag{138}$$

Now restrict the basis to obey the following suggestive relations

$$\theta_a \theta_b = -\theta_b \theta_a \tag{139}$$

$$\theta_a^2 = 0. \tag{140}$$

Then the new vector space has dimension $\dim(\mathcal{G}(V)) = 2^n$, and is the infinite dimensional span modulo the elements of the underlying vector space

$$\mathcal{G}(V) = \mathcal{S}^{\infty}(V)/v. \tag{141}$$

(**Exercise**) Check that this structure is well-defined. Note that this is exactly the space of differential forms for a tangent space V (also known as the space of "classical fermions").

The basis of $\mathcal{G}(V)$ is now

$$\{1,\, \theta_a,\, \theta_a\theta_b,\, \theta_a\theta_b\theta_c, \ldots\} \tag{142}$$

With $a = 1, \ldots, n$, followed by $1 \leq a < b \leq n$, $a < b < c$, etc.

Then a general element of the new vector space $f \in \mathcal{G}(V)$ is

$$f = \alpha + \sum_{p=1}^{n} \sum_{1 \leq j_1 < \cdots < j_p \leq n} \alpha_{j_1 j_2 \ldots j_p}\, \theta_{j_1} \theta_{j_2} \ldots \theta_{j_p} \tag{143}$$

$$\text{, where } \alpha_{j_1 j_2 \ldots j_p} \in \mathbb{C}. \tag{144}$$

Lecture 6: Grassmann Numbers

We left off with an object meant to be the classical version of a fermion: a Grassman number. It is the object of a vector space $\mathcal{G}_n(V)$, generated by basis $\{1, \theta_1, \ldots, \theta_n \in V\}$. Imposing the anticommutation relation

$$\{\theta_j, \theta_k\} = 0, \, \forall j, k = 1, \ldots, n \tag{145}$$

This is a 2^n-dimensional vector space with monomial basis

$$\{1, \theta_1, \theta_2, \ldots, \theta_n, \theta_1\theta_2, \ldots, \theta_{n-1}\theta_n, \ldots, \theta_1\theta_2 \ldots \theta_n\}. \tag{146}$$

An arbitrary element $f \in \mathcal{G}_n(V)$

$$f = f_0 + \sum_{j_1 < \cdots < j_p} f_p(j_1, \ldots, j_p) \, \theta_{j_1} \ldots \theta_{j_p} \tag{147}$$

We now define functions, linear and nonlinear, representations, complex numbers, calculus, derivatives, and integrals of Grassman numbers.

Functions of Grassman Numbers

Think of θ_j as anticommuting numbers/variables.

Note: A "wrong" definition is to let f be an infinitely differentiable function $f \in C^\infty(\mathbb{R})$, adjoin the symbol $f(\theta_j)$, and impose the anticommutation relation $\{f(\theta_j), \theta_k\} = 0, \, \forall \, j, k$. Then each f will produce a new anticommuting object, which leads to an uncountably infinite number of objects, and is not what we expect from a function.

Linear functions should be linear maps from the vector space to itself, $F \in \mathcal{M}_{2^n}(\mathbb{C})$, the space of $2^n \times 2^n$ matrices

$$F: \; \mathcal{G}_n(V) \to \mathcal{G}_n(V) \tag{148}$$

Nonlinear functions will be defined by analogy to the functional calculus of matrices; consider a matrix $M \in \mathcal{M}_m(\mathbb{C})$. As long as M is diagonalizable, define the nonlinear function $f(M) = S^{-1} f(D) S$, where M is diagonalized by S, such that $M = S^{-1} D S$. Since a diagonal matrix occupies a commutative algebra, we know how to define function for D and then rotate from D to M via S.

In summary so far, one strategy to define functions of Grassman numbers is to represent $\mathcal{G}_n(V)$ as matrices and use functional calculus. Working with representations, we now need matrices to represent the Grassman numbers.

Representations of $\mathcal{G}_n(V)$

Consider a map from our vector space to a concrete space of matrices

$$\pi : \ \mathcal{G}_n(V) \to M_d(\mathbb{C} \tag{149}$$

Which must obey the anticommutation relation

$$\{\pi(\theta_j), \pi(\theta_k)\} = 0, \ \forall j, k. \tag{150}$$

Construct the "Jordan-Wigner representation", which will look very familiar. Begin with a Hilbert space $\mathcal{H} = \mathbb{C}^{2^n}$, with dimension 2^n, and Pauli operators

$$\sigma^+ = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{151}$$

$$\sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{152}$$

Recall that these Pauli operators obey the relations

$$\{\sigma^+, \sigma^z\} = 0 \tag{153}$$

$$(\sigma^+)^2 = 0. \tag{154}$$

Construct the representation of $\mathcal{G}_n(V)$

$$\pi(\theta_1) = \sigma_1^+ \otimes \mathbb{I}_2 \otimes \cdots \otimes \mathbb{I}_{n-1} \otimes \mathbb{I}_n \tag{155}$$

$$\pi(\theta_2) = \sigma_1^z \otimes \sigma_2^+ \otimes \cdots \otimes \mathbb{I}_{n-1} \otimes \mathbb{I}_n \tag{156}$$

$$\cdots \tag{157}$$

$$\pi(\theta_n) = \sigma_1^z \otimes \sigma_2^z \otimes \cdots \otimes \sigma_{n-1}^z \otimes \sigma_n^+ \tag{158}$$

Single variable representation ($n = 1$)

Let $F \in C^\infty(\mathbb{R}, \mathbb{R})$ and $\mathcal{G}_1(V) \simeq \{a + b\theta : \ a, b \in \mathbb{C}\}$. This should be consistent with the functional calculus of the representation $\pi(\cdot)$, such that $F(\pi(\cdot)) = S^{-1}F(D)S$.
To evaluate the function F, we write out the Taylor series, evaluate at $x = \theta$, and impose the defined relations (e.g., $\theta_j^2 = 0$, $\forall j$)

$$F(x) = \sum_{j=0}^{\infty} \frac{F^{(j)}(x = 0)x^j}{j!} \tag{159}$$

$$F(\theta) = \sum_{j=0}^{\infty} \frac{F^{(j)}(\theta = 0)\theta^j}{j!} \tag{160}$$

$$F(\theta) = F^{(0)}(\theta = 0) + F^{(1)}(\theta = 0) \cdot \theta \tag{161}$$

Example 1:

$$F(x) = \sin(x) \tag{162}$$

$$F(\theta) = (x - \frac{x^3}{3!} + \dots)\big|_{x=\theta} \tag{163}$$

$$F(\theta) = \theta \tag{164}$$

Example 2:

$$F(x) = x + x^3 \to F(\theta) = \theta \tag{165}$$

Multiple variable representation

Let $F \in C^\infty(\mathbb{R}^n, \mathbb{R})$, with the Taylor expansion

$$F(\theta_1, \dots, \theta_n) = F^{(0)}(\theta_1 = 0, \dots, \theta_n = 0) + \sum_{j=1}^{n} \theta_j \frac{\partial F(0, \dots, 0)}{\partial \theta_j} + \mathcal{O}(\theta^2) \tag{166}$$

Example 1: $n = 2$

$$F(x, y) = e^{-\lambda xy} \tag{167}$$

$$F(\theta_1, \theta_2) = (1 - \lambda xy + \lambda^2 x^2 y^2 + \dots)\big|_{x=\theta_1, y=\theta_2} \tag{168}$$

$$F(\theta_1, \theta_2) = 1 - \lambda \theta_1 \theta_2 \tag{169}$$

Example 2: $n = 2$

$$F(x, y) = e^{-\lambda_1 x - \lambda_2 y} \tag{170}$$

$$F(\theta_1, \theta_2) = (1 - \lambda_1 \theta_1)(1 - \lambda_2 \theta_2) \tag{171}$$

$$F(\theta_1, \theta_2) = 1 - \lambda_1 \theta_1 - \lambda_2 \theta_2 + \lambda_1 \lambda_2 \theta_1 \theta_2 \tag{172}$$

Note that even though we lose higher order terms such as θ_j^2, nonlinear features are preserved in the multivariable case.

Complex Grassman Numbers

Let θ_1, $\theta_2 \in \mathcal{G}_2(V)$, a 4-dimensional vector space with basis $\{1, \theta_1, \theta_2, \theta_1\theta_2\}$, and define the quantities θ and $\theta^* \in \mathcal{G}_2(V)$ as

$$\theta = \frac{\theta_1 + i\theta_2}{\sqrt{2}} \tag{173}$$

$$\theta^* = \frac{\theta_1 - i\theta_2}{\sqrt{2}}. \tag{174}$$

To extend to the multivariable case, define

$$\theta_j^* = \frac{\theta_{j_1} - i\theta_{j_2}}{\sqrt{2}} \tag{175}$$

Where θ_{j_1}, $\theta_{j_2} \in \mathcal{G}_{2n}(V)$.

Grassman Derivatives

Define the derivative with respect to Grassman variables as the map

$$\partial_{\theta_j} : \ \mathcal{G}_n(V) \to \mathcal{G}_n(V) \tag{176}$$

Which obeys the relations

1. $\partial_{\theta_j}(\theta_k) = \delta_{jk}$
2. $\partial_{\theta_j}(\theta_{k_1} \dots \theta_{k_p}) = \delta_{jk_1}\theta_{k_2} \dots \theta_{k_p} - \delta_{jk_2}\theta_{k_1}\theta_{k_3} \dots \theta_{k_p} + \dots + (-1)^{p-1}\delta_{jk_p}\theta_{k_1} \dots \theta_{k_{p-1}}$.

The second relation follows from the anticommutation relation, and can be thought of as bringing the corresponding θ_j to the front via anticommutations and then differentiating.
For example,

$$\partial_{\theta_2}(\theta_1\theta_2) = -\partial_{\theta_2}(\theta_2\theta_1) = -\theta_1. \tag{177}$$

(**Exercise**) These relations can be extended by linearity for any Grassman number. This definition of ∂_{θ_j} obeys the product rule and the chain rule.

Grassman Integrals

Following the analogy of the common definite integral, the Grassman integral should be a linear map which obeys shift invariance, such that $\theta \to \theta + \eta$,

$$\int d\theta : \ \mathcal{G}_n(V) \to \mathbb{C}. \tag{178}$$

The only consistent definition to satisfy these two constraints for a single Grassman variable is

$$\int d\theta \ (a + b\theta) = b, \ \ a, b \in \mathbb{C} \tag{179}$$

Note the very interesting property here that, by this definition, the integral and the derivative are the exact same thing

$$\int d\theta \, (a + b\theta) = \partial_\theta(a + b\theta). \tag{180}$$

This also holds for the multivariable case, where the highest order term is picked off from the Taylor series

$$\int d\theta_n \dots d\theta_1 \, (f_0 + \sum_{j_1 < \dots < j_p} f_p(j_1, \dots, j_p)\theta_{j_1} \dots \theta_{j_p}) = f_n(1, \dots, n). \tag{181}$$

This is a weird definition, but it works, behaves correctly under change of variables, and does what we need it to do.

Example: $n = 2$ independent Grassman numbers

$$\int d\theta^* d\theta \, e^{-\lambda\theta^*\theta} = \int d\theta^* d\theta \, (1 - \lambda\theta^*\theta) \tag{182}$$

$$= \partial_{\theta^*}\partial_\theta \, (1 - \lambda\theta^*\theta) \tag{183}$$

$$\int d\theta^* d\theta \, e^{-\lambda\theta^*\theta} = \lambda \tag{184}$$

Note that the order of integration (θ first, θ^* second) is by convention.

Multivariable Gaussian Integrals

Consider the multivariable Gaussian integral

$$I = \int d\theta_1^* d\theta_1 \dots d\theta_n^* d\theta_n \; e^{-\sum_{j,k} \theta_j^* B_{jk} \theta_k} \tag{185}$$

Where $B^\dagger = B$ and we diagonalize via $U^\dagger B U = D$.

Make a change of variables $\theta'_j = U_{jk}\theta_k$, where U is a unitary matrix, such that the product of these new Grassman variables is (**Exercise**)

$$\theta'_1\theta'_2 \dots \theta'_n = \sum_{k_1,\dots,k_n} U_{1k_1}U_{2k_2} \dots U_{nk_n}\theta_{k_1}\theta_{k_2} \dots \theta_{k_n} \tag{186}$$

$$= \sum_{\pi \in S_n} U_{1\pi(1)} \dots U_{n\pi(n)} \mathrm{sgn}(\pi)\theta_1\theta_2 \dots \theta_n \tag{187}$$

$$\theta'_1\theta'_2 \dots \theta'_n = \det(U)\theta_1\theta_2 \dots \theta_n \tag{188}$$

So the Gaussian integral becomes

$$I = \int d\theta'^*_1 d\theta'_1 \dots d\theta'^*_n d\theta'_n \; e^{-\sum_{j,k} \theta'^*_j (UBU^\dagger)_{jk} \theta'_k} \det(U)\det(U^\dagger) \tag{189}$$

$$= \int d\theta'^*_1 d\theta'_1 \dots d\theta'^*_n d\theta'_n \; e^{-\lambda_1 \theta'^*_1 \theta'_1} \dots e^{-\lambda_n \theta'^*_n \theta'_n} \cdot (1) \cdot (1) \tag{190}$$

$$= \lambda_1\lambda_2 \dots \lambda_n \tag{191}$$

$$I = \det(B) \tag{192}$$

Recall that for the normal Gaussian integral case, we got $I = (\det(B))^{-1}$.

(**Exercise**) The generating functional for the Grassman calculus is, where J is a vector of Grassman numbers,

$$Z[J] = \int d\theta_1^* d\theta_1 \dots d\theta_n^* d\theta_n \; e^{-\theta^\dagger B\theta + J^\dagger\theta + \theta^\dagger J} \tag{193}$$

$$= e^{J^\dagger B^{-1} J}. \tag{194}$$

The moments are (**Exercise**)

$$\int d\theta_1^* d\theta_1 \dots d\theta_n^* d\theta_n \; \theta_j \theta_k^* e^{-\theta^\dagger B\theta} = \det(B) \cdot [B^{-1}]_{jk} \tag{195}$$

Side: The mixing of regular and Grassman numbers provides the basis for the supersymmetric method. Consider the Gaussian integral

$$\int d\Phi e^{-\Phi^\dagger M\Phi} = \int dx_1 \dots dx_n dx_1^* \dots dx_n^* \int d\theta_1^* d\theta_1 \dots d\theta_n^* d\theta_n \, e^{-\Phi^\dagger M\Phi} \tag{196}$$

$$= (\det A)(\det A)^{-1} \tag{197}$$

$$\int d\Phi e^{-\Phi^\dagger M\Phi} = 1 \tag{198}$$

Where $\Phi = (x_1, \dots, x_n, \theta_1, \dots, \theta_n)$ and M is a $2n \times 2n$ matrix

$$M = \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} \tag{199}$$

Lecture 7: Functional Quantization of the Dirac Field

We now employ Grassmann numbers/variables to build a path integral-like object that provides the n-point correlation functions for the Dirac (spinor) field.

Consider the Grassmann integral of the complex Grassmann variables θ and θ^*, the Grassmann Gaussian generating function

$$Z[J] = \left(\prod_{j=1}^{n} \int d\theta_j^* d\theta_j \right) e^{-\sum_{j,k} \theta_j^* B_{jk} \theta_k + \sum_j (J_j^* \theta_j + \theta_j^* J_j)} \tag{200}$$

Where the Grassmann variables and auxiliary fields J and J^* obey the anticommutation relations

$$\{\theta_j, \theta_k^*\} = \{\theta_j, \theta_k\} = \{\theta_j^*, \theta_k^*\} = 0 \tag{201}$$

$$\{J_j, \theta_k\} = \{J_j^*, \theta_k\} = \{J_j, J_k^*\} = 0. \tag{202}$$

Calculating these Gaussian integrals, the generating functional becomes

$$Z[J] = e^{-\sum_{j,k} J_j^* [B^{-1}]_{jk} J_k}. \tag{203}$$

Since the matrix B is unitary, such that $B^\dagger = B$, the generating functional $Z[J]$ is Hermitian.

Note that the generating functional takes a vector of Grassmann numbers, anticommuting objects, as input and yields an expression quadratic in the Grassmann numbers which evaluates to a real number, a commuting object, since observables correspond to Hermitian operators and real numbers as their eigenvalues; the expectation value must always be a real number, not a Grassmann number.

Recall the Dirac spinor field which is what we mean to be classical fermions. The classical fermion is represented by the 4-component spacetime vector

$$\psi(x) \to M(\Lambda)\psi(\Lambda^{-1}x) \tag{204}$$

With the representation of the Lorentz group

$$M = e^{-\frac{i}{2}\omega_{\mu\nu} S^{\mu\nu}}. \tag{205}$$

Where $S^{\mu\nu} = \frac{i}{4}[\gamma^\mu, \gamma^\nu]$ and $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}$. An object that transforms according to the transformation law with this representation we call a Dirac spinor. The representation, and thus the generators, of the Poincaré group

easily follows.

The Hamiltonian, or generator of time translations, that follows from this spinor object solves the Dirac equation $(i\partial_\mu\gamma^\mu - m)\psi = 0$. Recall that we defined the conjugate-like object $\bar{\psi} = \psi^\dagger\gamma^0$ to induce Lorentz invariance for the Lagrangian density $\mathcal{L} = \bar{\psi}(-i\not{\partial} - m)\psi$, where the slash notation denotes $\not{A} = A_\mu\gamma^\mu$.

Thus far, we've been thinking about $\psi = (\psi_1, \psi_2, \psi_3, \psi_4)$ as the classical spinor-valued Dirac field, but it is really the single-particle component that is needed to build the classical Dirac field with anticommuting objects, since we guess to employ the anticommutation relation of quantum Dirac field operators

$$\{\hat{\psi}(x), \hat{\psi}^\dagger(y)\} = \delta^{(4)}(x - y). \tag{206}$$

So, to build a quantum theory, recall that we, in the case of creation and annihilation operators, for example,

1. Pick a classical single-particle theory
2. Put hats on the field operators to quantize them
3. Make an algebra for the quantized field operators to obey
4. Find representations of that algebra.

Alternatively, we can find some classical Dirac field that we quantize via the path integral, and use the path integral as a tool to guess the quantum theory.

Guess 1 (Wrong):

Let the classical field $\psi(x)$ consist of real numbers. Then the corresponding path integral $\int \mathcal{D}\psi\mathcal{D}\bar{\psi}\, e^{iS}$ will not yield the n-point correlation functions for the quantum Dirac (fermionic) field, but will yield the n-point correlation functions for the bosonic field.

Guess 2 (Correct):

Let the classical field $\psi(x)$ consist of Grassmann numbers, a Grassmann-valued field as the classical Dirac field. Then the path integral will yield the quantum Dirac field.

A Grassmann-valued field (4-vectors) can be understood via sheaf theory and ringed spaces of Grassmann numbers on manifolds.

An alternative way to make sense of a Grassman-valued field is by discretization to a lattice of spacing ϵ, compactified to a torus. Consider a map from spacetime to the 4-dimensional torus

$$\mathbb{R}^{1,3} \to (\mathbb{Z}/N\mathbb{Z})^{\otimes 4} \tag{207}$$

Consider the (0+1)-dimensional case, mapping continuous spacetime coordinates to discrete coordinates: $x_j \to \epsilon j$, where $j \in \mathbb{Z}/N\mathbb{Z}$, and discretize the Grassmann numbers to the lattice to define the classical Dirac field

$$\psi_j \equiv \psi(x_j) \equiv \psi(\epsilon j) \tag{208}$$

$$\psi_j^\dagger \equiv \psi^\dagger(x_j) \equiv \psi^\dagger(\epsilon j). \tag{209}$$

So, the (discrete) classical Dirac field is a list of $8 \cdot N^4$ Grassmann numbers, since N^4 is the number of lattice sites and each of the two field "operator" contains 4 components.

Note that if we work in momentum space, the Fourier coefficients will be made to be Grassmann numbers.

The continuous classical Dirac field comes from the limit of the lattice spacing vanishing $\epsilon \to 0$, the number of sites tending to infinity $N \to \infty$, and the size of the torus tending to infinity $L = N\epsilon \to \infty$.

Now, to build the quantum theory corresponding to these classical objects via the path integral formalism, we require an action, beginning with the discretization of the Lagrangian density $\mathcal{L} = \bar{\psi}(i\partial\!\!\!/ - m)\psi$. For the Dirac field, the discretized Lagrangian density is

$$\mathcal{L}(\psi_j, \bar{\psi}_j) = \sum_{j \in (\mathbb{Z}/N\mathbb{Z})^{\otimes 4}} i\bar{\psi}_j \left(\gamma^\mu \left(\frac{\psi_{j+\hat{\mu}} - \psi_j}{\epsilon} \right) \right) - m\bar{\psi}_j \psi_j. \tag{210}$$

Where ψ_j and $\bar{\psi}_j$ are $4D$ spinors of Grassmann numbers, we've employed the forward-difference to represent the partial derivative, and $\hat{\mu}$ is a unit vector in the μ^{th} direction.

With the Lagrangian density, we can calculate the action $S = i \int_{-T}^{T} dt\, \mathcal{L}(\psi_j, \bar{\psi}_j)$ and the n-point correlation functions for the Grassmann-valued quantum field operators.

For example, for the Grassmann variables, define the 2-point correlation function

$$\langle 0 | \mathcal{T}[\hat{\psi}(x), \hat{\bar{\psi}}(y)] | 0 \rangle \equiv \lim_{T \to \infty(1-i\epsilon)} \frac{\int \mathcal{D}\bar{\psi}\mathcal{D}\psi \ \psi(x)\bar{\psi}(y) e^{iS}}{\int \mathcal{D}\bar{\psi}\mathcal{D}\psi \ e^{iS}} \tag{211}$$

To calculate the 2-point correlation function, we continue to follow the prescription

1. Discretize the field and the action
2. Evaluate the path integral

3. Take the continuous limit

Evaluate the path integral to find that the 2-point function is

$$\langle 0|\,\mathcal{T}[\hat{\psi}(x),\hat{\bar{\psi}}(y)]\,|0\rangle = S_F(x-y) = \int \frac{d^4k}{(2\pi)^4}\,\frac{ie^{-ik\cdot(x-y)}}{\not{k}-m+i\epsilon} \tag{212}$$

Side note (topic of ongiong research): Fermion doubling is a topological artifact of incorrectly placing fermions on a lattice and taking the continuous limit. Extra fermions, called doublers, appear in the calculation, as the dispersion relation $\omega(k)$ becomes nonlinear and crosses the k-axis more than once. The expected dispersion relation is linear $\omega(k) = ak,\ a > 0$. In discretization, we must accept this effect and learn how to work around it. This is done for conveience, since without discretization, evaluating the 2-point function requires many more tricks.

Generating Functional for the Dirac Field

Define the generating functional for the Dirac field in terms of two independent Grassmann-valued functions

$$Z[J(x),\bar{J}(y)] \equiv \int \mathcal{D}\bar{\psi}\mathcal{D}\psi\ e^{i\int d^4x\ (\bar{\psi}(i\not{\partial}-m)\psi+\bar{J}\psi+\bar{\psi}J)}. \tag{213}$$

Where J and $\bar{J}$ are Grassmann-valued (auxiliary) source fields that will be set to zero after differentiation. Calculating the generating functional will yield all n-point functions via functional derivatives, made possible by the employment of Grassmann numbers and functional quantization versus canonical quantization. By completing the square and simplifying the expression for the generating functional we get (**Exercise**)

$$Z[J,\bar{J}] = Z_0 e^{-\int d^4x d^4y\ \bar{J}(x)S_F(x-y)J(y)} \tag{214}$$

Where $Z_0 = Z[J=0,\bar{J}=0]$. Recall that for Grassmann numbers, the rules of differentiation include sign-switching and go like

$$\frac{d}{d\eta}\theta\eta = -\theta\frac{d}{d\eta}\eta = -\theta \tag{215}$$

So the n-point correlation function is then

$$\langle 0|\,\mathcal{T}[\psi^{(\alpha_1)}(x_1)\dots\psi^{(\alpha_n)}(x_n)]\,|0\rangle = Z_0^{-1}\left(i(-1)^{\alpha_1+1}\frac{\delta}{\delta J^{\alpha_1}}\right)\dots\left(i(-1)^{\alpha_n+1}\frac{\delta}{\delta J^{\alpha_n}}\right)Z[J,\bar{J}] \tag{216}$$

Where $\psi^{(\alpha)}(x) = \psi(x)$ for $\alpha = 0$ and $\psi^{(\alpha)}(x) = \bar{\psi}(x)$ for $\alpha = 1$.

Check (**Exercise**) that the quantum 2-point correlation function comes out to be the expected

$$\langle 0|\,\mathcal{T}[\hat{\psi}(x)\hat{\bar{\psi}}(y)]\,|0\rangle = S_F(x-y). \tag{217}$$

Interactions of Fermions and Bosons

The path integral is a great tool for guessing Feynman rules as well, since we can expand in a Taylor series and recognize patterns that represent certain symmetries and diagrams. WIthout needing to introduce too much gauge theory, we introduce *massive quantum electrodynamics (QED)*, a quantum field theory that models the interaction of fermions and (massive) bosons. We expect the photon (boson) mass to be zero, but consider it massive for now, and note that the upper bounds on the mass of the photon have been calculated to be nonzero (10^{-20}).

Consider the Lagrangian density

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu(\partial_\mu - ieA_\mu) - m_f)\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}m_b^2 A_\mu A^\mu. \tag{218}$$

There are six fields represented here: the fermion fields ψ and $\bar{\psi}$, the 4 boson fields A_μ, and the tensor $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$.

Following the path integral quantization, the classical 2-point correlation function is

$$\langle 0|\,\mathcal{T}[\psi(x)\bar{\psi}(y)]\,|0\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\int \mathcal{D}\psi\mathcal{D}\bar{\psi}\mathcal{D}A\ \psi(x)\bar{\psi}(y)e^{iS}}{\int \mathcal{D}\psi\mathcal{D}\bar{\psi}\mathcal{D}A\ e^{iS}}. \tag{219}$$

Write the action in terms of the free theory and the interacting theory

$$\begin{aligned} S &= S_0 + S_{int} \\ S &= \left(\int d^4x\,(\bar{\psi}(i\partial\!\!\!/ - m_f)\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}m_b^2 A_\mu A^\mu\right) \\ &\quad + \left(-ie\int d^4x\,\bar{\psi}A_\mu\gamma^\mu\psi\right). \end{aligned}$$

Then the quantized 2-point correlation function for massive QED with the Taylor expansion is an infinite series of n-point correlation functions for the Grassmann-valued Gaussian path integrals

$$\langle 0|\,\mathcal{T}[\hat{\psi}(x)\hat{\bar{\psi}}(y)]\,|0\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\int \mathcal{D}\psi\mathcal{D}\bar{\psi}\mathcal{D}A\ e^{iS_0}\psi(x)\bar{\psi}(y)(1 - ie\int d^4x\,\bar{\psi}A_\mu\gamma^\mu\psi + \dots)}{\int \mathcal{D}\psi\mathcal{D}\bar{\psi}\mathcal{D}A\ e^{i(S_0+S_{int})}}. \tag{220}$$

The Feynman rules for massive QED, which can be deduced via patterns from the Taylor series, are

1. Draw a straight line from a to b with momenta p for each fermion and associate $\left(\frac{i}{\not{p}-m_f+i\epsilon}\right)_{ab}$

2. Draw squiggly line from α to β with momenta q for each boson and associate $\left(\frac{-i}{k^2-m_b^2+i\epsilon}\right)\delta_{\alpha\beta}$
3. To each vertex associate $ie\gamma^\mu$
4. Enforce momentum conservation at vertices: $(2\pi)^4\delta^{(4)}(\Sigma_{in}p - \Sigma_{out}q)$
5. Integrate over undetermined momenta
6. Amputate external lines
7. Incoming fermions $\eta^a(p)$ and outgoing fermions $\bar{\eta}^b(p)$
8. (-1) for each closed fermion loop.

This prescription creates infinities, and we will visit renormalization to tame these infinities in the following lectures.

Lecture 8: Renormalization

This is an incredibly important lecture where we ask, "How do we do physics?" and "How do we make progress in physics?"

1. Observation
   → Empirical data
2. Explanation
   → Data compression
3. Understanding
   → Models
   E.g., Hamiltonian with Hilbert space, neural network, list of data
4. Prediction
   → New observation
5. Repeat from Step 1

This algorithm yields an increasingly smaller list of plausible models that correspond with the observed data. To the physicist, a model is often a Hamiltonian (or Lagrangian), with an associated Hilbert space $\mathcal{H}_\Lambda$, that depends on some list of unknown parameters $z_j \in \mathbb{R}$, for $j = 1, \ldots, n$

$$\hat{H} \to \hat{H}_\Lambda = \hat{H}_\Lambda(z_1, \ldots, z_n) \tag{221}$$

Where Λ is a list of the degrees of freedom that we wish to explain. The degrees of freedom may be finite or infinite, and hopefully there are tricks to tame the infinite ones.

Some additional observations

When we speak of empirical data, we mean expectation values of Hermitian operators $\langle A_j \rangle = \alpha_j$, which can be precisely defined such that $\alpha_j = \delta_{\alpha_j}$ or can have a spread of uncertainty.

For each model $\hat{H}_\Lambda(z_1, \ldots, z_n)$ that we make a prediction for $\langle A_j \rangle$, if the parameters z_j do not yield the correct expectation value, then we reject that set of parameters for the model, and end up with a map

$$\langle A_j \rangle = f_j(z_1, \ldots, z_n; \Lambda). \tag{222}$$

This map is the *exact solution* or our prediction for that model. Note that this map is many-to-one, and is, thus, not invertible; there are many sets of parameters that can yield the same expectation value.

We say that a model $\hat{H}_\Lambda(z_1, \ldots, z_n)$ is "simpler" than another model $\hat{H}'_\Lambda(z'_1, \ldots, z'_{n'})$ if one or both of the following conditions are satisfied: $n < n'$ and/or $|\Lambda| > |\Lambda'|$.

The parameters $z_1, \ldots, z_n$ are essentially coupling constants, and are not directly observable and not operationally well-defined.

All predictions $f_j = \langle A_j \rangle$ must be finite and real.

It is possible to give finite predictions in terms of infinite parameters.

Renormalization and QFT

In the context of quantum field theory, consider the interacting ϕ^4 theory with Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4 \tag{223}$$

We have spent many weeks approximately computing the map $f_j(m, \lambda)$, for fixed m and λ, and have encountered infinity many times already in these calculations.

The degrees of freedom that we wish to explain in this context Λ are the *momentum modes* (of interacting bosons).

To tame these infinities, we first *impose a cutoff* $|\Lambda| < \infty$, where Λ is an arbitrary parameter. So, predictions will change with respect to the chosen cutoff, since $\langle A_j \rangle$ depends on Λ, such that $\langle A_j \rangle = f_j(z_1, \ldots, z_n; \Lambda)$.

We can declare victory if we can invert the prediction $f_j(z_1, \ldots, z_n; \Lambda)$ and move the Λ-dependence onto the parameters, such that $z_j = z_j(\Lambda)$.

A theory which allows $\langle A_j \rangle = f_j(z_1(\Lambda), \ldots, z_n(\Lambda); \Lambda)$, $\forall \Lambda$ and fixed n, is called *renormalizable*.

Side note about parameters: Note that the mass of an electron is the measured value, but in the model it is defined by the imposed cutoff. For a different cutoff, the coupling constant m may be different than the actual mass of the electron, but in the "correct", or "most correct", model, we call it the "mass of the electron".

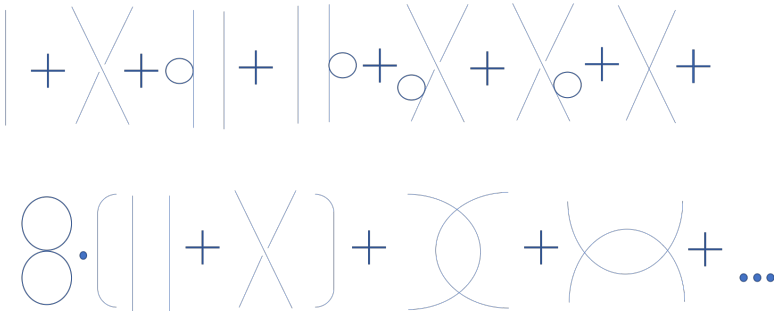
Scattering Amplitude in ϕ^4 Theory

In ϕ^4 theory, we focus on one prediction in particular, the scattering amplitude.

Note on things to come: the combinatorial proof that ϕ^4, and other models, is renormalizable.

Recall that the scattering S-matrix for ϕ^4 theory with no cutoff, or $|\Lambda| = \infty$, blows up to infinity

$$\langle p_1 p_2 | \, S \, | p_A p_B \rangle = \tag{224}$$



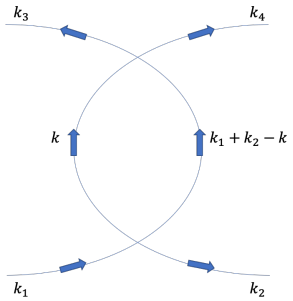
We can eliminate most of the infinite terms/diagrams by redefining the ground state energy E_0 and rest mass of the electron m_0 and rescaling λ, such that we dont have to set it to zero to make the theory work.

The first term/diagram that rescaling λ does not work for introduces a logarithmic divergence and has the form

$$I = \frac{(-i\lambda)^2}{2} \int \frac{d^4k}{(2\pi)^4} \, \frac{i}{k^2 - m^2 + i\epsilon} \frac{i}{(k_1 + k_2 - k)^2 - m^2 + i\epsilon}. \tag{225}$$

Therefore, we must impose a cutoff $|\Lambda| \neq \infty$ in order to calculate the S-matrix for ϕ^4 theory.

Impose a cutoff on the momenta $|k| < k_c \in \mathbb{R}$. Then the integral above becomes (**Exercise**)



$$I = \frac{(-i\lambda)^2 i^2}{2} \int_\Lambda \frac{d^4k}{(2\pi)^4} \; \frac{i}{k^2 - m^2 + i\epsilon} \frac{i}{(k_1 + k_2 - k)^2 - m^2 + i\epsilon} \tag{226}$$

$$I = 2iC \log \frac{k_c^2}{(k_1 + k_2)^2}. \tag{227}$$

Then the scattering amplitude to order $\mathcal{O}(\lambda^2)$ is

$$\mathcal{M} = \mathcal{M}(k_c) = -i\lambda + iC\lambda^2 \left(\log \left(\frac{k_c^2}{(k_1 + k_2)^2} \right) + \log \left(\frac{k_c^2}{(k_1 - k_3)^2} \right) + \log \left(\frac{k_c^2}{(k_1 - k_4)^2} \right) \right). \tag{228}$$

The parameter $z_2 = \lambda$ can be fit to the experiment and model, as we do not accept that it is fixed "at the beginning of the Universe", and we can declare victory by allowing $z_2 = z_2(k_c) = \lambda(k_c)$.

To solve for $\lambda(k_c)$, let the scattering amplitude be the experimental value $\mathcal{M}(k_c, \lambda) = \mathcal{M}_{exp}$ and solve the differential equation that allows λ to vary with respect to k_c and match up to $\mathcal{M}_{exp}$

$$k_c \frac{d\lambda}{dk_c} = 6C\lambda^2 + \mathcal{O}(\lambda^3). \tag{229}$$

Lecture 9: Renormalizability (of ϕ^4 Theory)

This topic covers results of Bogoliubov, Parasiuk, Hepp, Zimmermann, or the BPHZ renormalization scheme.

A *renormalizable* theory is a cutoff field theory determined by a finite number of parameters $\hat{H}_\Lambda = \hat{H}(z_1, \ldots, z_n; \Lambda)$ such that all observables $\hat{A}_\alpha$, the expectation value of $\hat{A}$ can be matched to experimentally determinable quantities for any choice of cutoff Λ by redefining the parameters $z_j = z_j(\Lambda)$

$$\langle \hat{A}_\alpha \rangle_{z_1,\ldots,z_n,\Lambda} = \langle \hat{A}_\alpha \rangle_{z_1(\Lambda),\ldots,z_n(\Lambda)} = \text{Experiment}(\alpha). \tag{230}$$

There is a weak form of renormailzability that retains dependence of the expectation value on the cutoff

$$\langle \hat{A}_\alpha \rangle_{z_1(\Lambda),\ldots,z_n(\Lambda)} = \text{Experiment}(\alpha) + \mathcal{O}\left(\frac{1}{k_c}\right). \tag{231}$$

This dependence can be worked around, since as $k_c \to \infty$, the inverse goes to zero.

In the ϕ^4 interaction, we have three parameters with the Hamiltonian

$$\hat{H}(m, \lambda, z; \Lambda) \tag{232}$$

Where z is the field strength renormalization parameter.

Is ϕ^4 theory, by this definition, renormalizable?

If this is true, then we are allowed to fit an infinite number of quantities to experimentally determinable quantities by fitting only three parameters: m, λ, and z. Very cool!

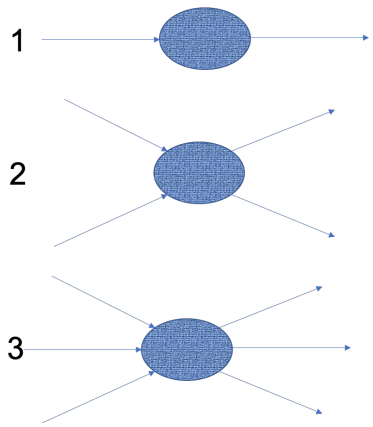
Degree of Divergence

Consider a diagram with B_E external lines. The diagram has a *superficial degree of divergence* D if it diverges with the cutoff as k_c^D. For $D = 0$, we say that the diagram has logarithmic divergence: $\log(k_c)$.

Theorem: The degree of divergence is equal to the number of spacetime dimensions minus the number of external lines

$$D = 4 - B_E. \tag{233}$$

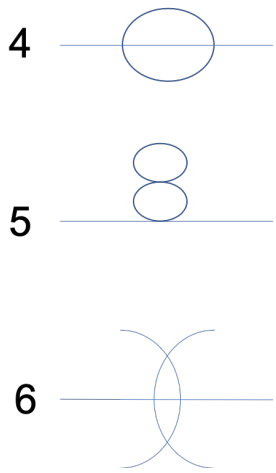
Examples:



1. $B_E = 2 \implies D = 2 \implies \sim k_c^2$
2. $B_E = 4 \implies D = 0 \implies \sim \log(k_c)$
3. $B_E = 6 \implies D = -2 \implies \sim \frac{1}{k_c^2}$

Note that as B_E and k_c increase, the divergences become increasingly less observable. Each pair of incoming/outgoing particles contributes a propagator proportional to k_c^{-2} to the $4D$ momentum integral $\sim \int \frac{d^4k}{(2\pi)^4} \left(\frac{i}{k^2 - m^2 + i\epsilon}\right)^{\frac{B_E}{2}}$.

Some more notation: B_I is the number of internal lines, V is the number of vertices, and L is the number of closed loops.



1. $B_I = 3$, $V = 2$, $L = 2$, $D = 2$
2. $B_I = 3$, $V = 2$, $L = 2$, $D = 2$

3. $B_I = 5$, $V = 4$, $L = 2$, $D = -2$

Proof:

The number of loops corresponds directly to the number of undetermined momenta ($4D$ momentum space integrals). Morally, we say that $\int \frac{d^4k}{(2\pi)^4} \sim k_c^4$.

It seems that there are B_I such integrals, but momentum conservation reduces the total number of loop integrals to

$$L = B_I - (V - 1). \tag{234}$$

Each vertex has four lines and each line connects two vertices, such that

$$4V = B_E + 2B_I. \tag{235}$$

Now, recall that for each loop there is a factor $\sim k_c^4$ from the integral, and for each line there is a factor k_c^{-2} from the propagator. Then

$$D = 4L - 2B_I = 4 - B_E. \tag{236}$$

Exercise: Prove this result for n-dimensional spacetime.

Physical or Renormalized Perturbation Theory

Consider the parameterized Lagrangian

$$\mathcal{L} = \mathcal{L}(z_1 = m, z_2 = \lambda, z_3 = z) = \frac{1}{2}(z^2(\partial_\mu\phi)^2 - z^2m^2\phi^2) - z^4\frac{\lambda}{4!}\phi^4. \tag{237}$$

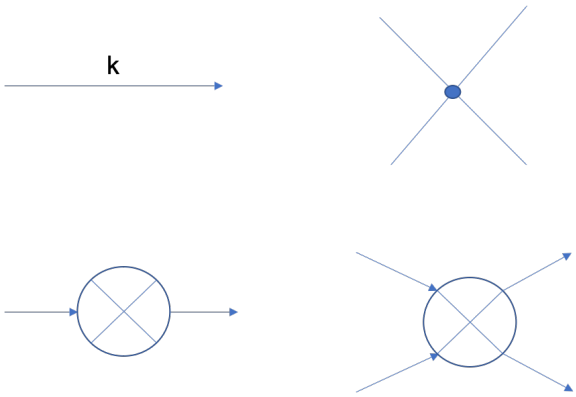
Rewrite this in terms of the physical Lagrangian, the one that has been successful in corresponding with experimental data, and counter terms dependent on three new parameters A, B, and C,

$$\mathcal{L} = \mathcal{L}_{\text{phys}} + \text{counter terms} \tag{238}$$

$$\mathcal{L} = \left(\frac{1}{2}(z^2(\partial_\mu\phi)^2 - m_{\text{phys}}^2\phi^2) - z^4\frac{\lambda_{\text{phys}}}{4!}\phi^4\right) + \left(A(\partial_\mu\phi)^2 + B\phi^2 + C\phi^4\right) \tag{239}$$

Now, think of these parameters as additional interactions that can be shifted to eliminate the dependence on the cutoff. These parameters are determined iteratively by the constraint that physically observable quantities do not depend of the momenta k.

The Feynman rules for the renormalized ϕ^4 theory are the same rules plus two more due to an additional type of vertex that depend of the "additional interaction" parameters.



Top-left:	$\frac{i}{k^2 - m^2_{\text{phys}} + i\epsilon}$
Top-right:	$-i\lambda_{\text{phys}}$
Bottom-left:	$2i(Ak^2 + B)$
Bottom-right:	$4! \cdot iC$

So, counter terms are added to the Lagrangian as "additional interactions", which introduce new Feynman diagrams. The parameters are determined iteratively to order λ^N_{phys}, at which we call them A_N, B_N, and C_N, and nothing depends on the cutoff.

The next iteration $N + 1$ is determined by requiring that the new propagator (bottom-left diagram above) at $\lambda^{N+1}_{\text{phys}}$ has a pole at m_{phys} with residue equal to one. We also require that the scattering amplitude to order $\mathcal{O}(\lambda^{N+1}_{\text{phys}})$ is equal to $-i\lambda_{\text{phys}}$ (bottom-right diagram above).

Non-Renormalizable Theories

A non-renormalizable theory requires an infinite number of parameters to ensure that operationally-defined quantities do not depend on the cutoff.

How do these theories appear in renormalized perturbation theory?

For instance, consider the Lagrangian

$$\mathcal{L} = \mathcal{L}^0_{\text{phys}} + \mathcal{L}^{int}_{\text{phys}}(\lambda) + (\text{counter terms}). \tag{240}$$

Calculate the scattering amplitude to order $\mathcal{O}(\lambda^N_{\text{phys}})$, and we will see that we need counter terms to eliminate dependence on the cutoff k_c, but as we go to higher and higher order in λ_{phys}, we need more and more counter terms, and this will continue and diverge, requiring an infinite number of counter terms and associated parameters to eliminate the cutoff dependence.

Lecture 10: Abelian Gauge Theory (Quantum Electrodynamics)

Why study abelian gauge theory?
An example of an abelian gauge theory is quantum electrodynamics and the electromagnetic field, as well as $SU(2)$, and $SU(3)$ gauge bosons.

In developing a gauge theory, we will follow the same route of specifying symmetries, giving rise to invariants, of the theory and then, via quantization, look for (projective) unitary representations of the group which are local.

As opposed to other field theories, the gauge theory should be symmetric under a local gauge group $\mathcal{G}$ which acts independently at each location in spacetime.

For example, consider the circle group $U(1)$, which consists of all complex numbers with absolute value equal to 1 under multiplication, the roots of unity

$$U(1) = \{e^{i\theta} : \theta \in [0, 2\pi)\}. \tag{241}$$

This symmetry group gives rise to the local gauge group $\mathcal{G}$ as the group of transformations from $1+3$-dimensional space time, Minkowski space, to the circle group

$$\mathcal{G} = \{g : \mathcal{M}_{1+3} \to U(1)\} \tag{242}$$

How does the local gauge group act on fields?

Consider the DIrac field, where we want a $U(1)$ gauge-invariant quantum field theory of electrons. Elements of $\mathcal{G}$ act independently on each spacetime location$x \in \mathcal{M}_{1+3}$. Equivalently, there is a copy of $U(1)$ attached to each x acting independently of each other.

$$g : \psi(x) \to \pi(g(x))\psi(x) = e^{i\alpha(x)}\psi(x) \tag{243}$$

Where $\alpha(x) \in [0, 2\pi)$ is a phase factor.

Which theories are invariant under the Poincaré group and the local gauge group $\mathcal{G}$?

As it stands, we have an empty set of theories that are invariant under the local gauge group. Begin populating it by building a Lagrangian density, with a classical, continuous spacetime, and find which kinds of terms will be invariant.

Terms of the Lagrangian density for the Dirac field that we already know to be invariant are $\bar{\psi}\psi$ and $(\bar{\psi}\gamma^\mu\psi)^2$ (contracted with itself).

Note that the quantity $\bar{\psi}\not{\partial}\psi$ is not invariant, as it is not well-defined! Why?

The differential operator acting on the field $\partial_\mu\psi$ is a limit

$$\partial_\mu\psi \equiv \lim_{\epsilon\to\infty} \frac{\psi(x+\epsilon n^\mu)-\psi(x)}{\epsilon}. \tag{244}$$

And the local gauge group acting on this quantity results in oscillatory terms that do not converge as $\epsilon \to 0$ (e.g., $\sim e^{i\cdots}$). The limit does exist under gauge transformations, since

$$g(\partial_\mu\psi) = \lim_{\epsilon\to\infty} \frac{1}{\epsilon}(e^{-\alpha(x+\epsilon n^\mu)}\psi(x+\epsilon n^\mu) - e^{i\alpha(x)}\psi(x)). \tag{245}$$

This theory is boring at this moment, as the Lagrangian has only two gauge-invariant terms. To introduce dynamics and impose the desired symmetry, we need the derivative, but there is no way to do it with a single fermion (or boson) field. Therefore, we introduce an *auxiliary field* which transforms nontrivially on the local gauge group (c.f., adding a catalyst in chemistry opens new thermodynamic paths for reactions to take place, as the catalyst is not consumed, but is used, in the reaction to lowers the free energy of the reaction).

Introduce the *parallel transporter*, a recipe to compare a field at two independent spacetime locations, which is dependent on the path γ

$$U_\gamma(y,x) \in U(1),\ \forall\, x,y \in \mathcal{M}_{1+3} \tag{246}$$

To ensure that $U(y,x)$ is a gauge invariant comparator, and allow us to compare the two spacetime locations in a gauge invariant, we require that the local gauage group act as

$$g:\ U(y,x) \to e^{i\alpha(y)}U(y,x)e^{-i\alpha(x)}. \tag{247}$$

So, we define $\psi(x)$ to be *parallel transported* to y as

$$g:\ U(y,x)\psi(x) \to e^{i\alpha(y)}U(y,x)e^{-i\alpha(x)}\psi(x) \tag{248}$$

$$= e^{i\alpha(y)}U(y,x)\psi(x) \tag{249}$$

Does such an object $U(y,x)$ exist?

Yes, they exist and are made rigorous in the formalism of fibre bundles and principal bundles. Note that $U(y,x)$ is a not a field and is nonlocal object, but it is expressable in terms of local objects and local data.

So, $\psi(x+\epsilon n^\mu)$ and $U(x+\epsilon n^\mu,x)\psi(x)$ transform the same way under the local gauge group $\mathcal{G}$, and we can introduce dynamics to the theory and define the *covariant derivative* as

$$D_\mu \psi(x) \equiv \lim_{\epsilon \to 0} \frac{\psi(x + \epsilon n^\mu) - U(x + \epsilon n^\mu, x)\psi(x)}{\epsilon} \tag{250}$$

What about the parallel transporter $U(y, x)$?

Suppose that $U(x + \epsilon n^\mu, x)$ is continuous and differentiable near x, and apply the Taylor series expansion

$$\begin{aligned} U(x + \epsilon n^\mu, x) &= U(x, x) + \epsilon n^\mu \partial_\mu U(x, x) + \dots && (251) \\ &= 1 + \epsilon n^\mu \partial_\mu U(x, x) + \dots && (252) \\ &= 1 - i\epsilon n^\mu A_\mu(x) + \dots && (253) \end{aligned}$$

Where we tried $U(x, x) = 1$, since 1 transforms correctly under the gauge group, and it is traditional to call $\partial_\mu U(x, x) = -i\alpha A_\mu(x)$, where α is the fine structure constant. Now, $A_\mu(x)$ is not arbitrary, and must satisfy some constraints.

How does $A_\mu(x)$ transform under $\mathcal{G}$?

Apply a local gauge transformation $g \in \mathcal{G}$ to the last line above

$$e^{i\alpha(x+\epsilon n^\mu)} U(x + \epsilon n^\mu, x) e^{-i\alpha(x)} = 1 - ie\epsilon n^\mu A_\mu(x) + in^\mu \partial_\mu \alpha(x) + \dots \tag{254}$$

Thus, the auxiliary field transforms under the local gauge as

$$g: \; A_\mu(x) \to A_\mu(x) - \frac{1}{e} \partial_\mu \alpha(x) \tag{255}$$

And this gives us a definition of $A_\mu(x)$ in one gauge, and this transformation law allows us to change gauge, or basis. Finding the form of $A_\mu(x)$ that satisfies this local gauge transformation is tantamount to having an infinitesimal method for building the parallel transporter object.

Put this all together into the covariant derivative (**Exercise**)

$$D_\mu \psi(x) = \partial_\mu \psi(x) + i\alpha A_\mu(x) \psi(x) \tag{256}$$

Which is now a purely local object, dependent on only one spacetime location x.

Furthermore, the covariant derivative of the field transforms under the local gauge group by introducing the phase factor, the same as the action of $\mathcal{G}$ on the field ψ itself

$$g: \; D_\mu \psi(x) \to e^{i\alpha(x)} D_\mu \psi(x) \tag{257}$$

So, we've built the parallel transporter in terms of the local field $A_\mu(x)$, and we have a derivative object that transforms correctly under the local gauge group. Now, build the Lagrangian density

$$\mathcal{L} = \bar{\psi} \not{D} \psi - m \bar{\psi} \psi + \text{auxiliary field term(s)} \tag{258}$$

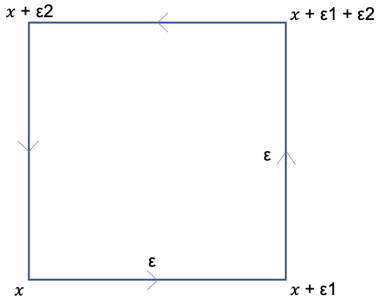
To include the auxiliary field $A_\mu(x)$ in the quantization we need to endow it with dynamics.

How do we give $A_\mu(x)$ dynamics?

To first order,

$$U(x + \epsilon n^\mu, x) \sim e^{-ie\epsilon n^\mu A_\mu(x + \frac{1}{2}\epsilon n^\mu)}. \tag{259}$$

Use the parallel transporter to build a *plaquette operator*, which transverses an object around a square of dimension ϵ



$$U_\square(x) = U(x, x + \epsilon\hat{1}) \cdot U(x + \epsilon\hat{1}, x + \epsilon(\hat{1} + \hat{2})) \cdot U(x + \epsilon(\hat{1} + \hat{2}), x + \epsilon\hat{2}) \cdot U(x + \epsilon\hat{2}, x). \tag{260}$$

The plaquette operator $U_\square$ is gauge invariant, and, working out the Taylor series, we can write it in terms of the auxiliary field (**Exercise**)

$$U_\square(x) = e^{-i\epsilon\alpha(-A_2(x + \frac{1}{2}\epsilon\hat{2}) - A_1(x + \frac{1}{2}\epsilon\hat{1} + \epsilon\hat{2}) + A_2(x + \epsilon\hat{1} + \frac{1}{2}\epsilon\hat{2}) + A_1(x + \frac{1}{2}\epsilon\hat{1})) + \mathcal{O}(\epsilon^3)}. \tag{261}$$

Expand in ϵ (**Exercise**)

$$U_\square(x) = 1 - i\epsilon^2 e(\partial_1 A_2 - \partial_2 A_1) + \mathcal{O}(\epsilon^3). \tag{262}$$

The choice of direction in this derivation is arbitrary, leaving 16 possible choices; construct the 2-tensor

$$F_{\mu\nu}(x) = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{263}$$

Note that $F_{\mu\nu}$ is locally gauge invariant, Lorentz invariant, but it is not Poincaré invariant with those spacetime indices. Construct the gauge, Lorentz, and Poincaré invariant object from the $F_{\mu\nu}$

$$-\frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \tag{264}$$

Then the first nontrivial Langrangian density we can construct is the exact Lagrangian for quantum electrodynamics: an electromagnetic field minimally coupled to the Dirac field.

$$\mathcal{L} = \bar{\psi}(\not{D} - m)\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \tag{265}$$

The term "minimally coupled" means that the theory is renormalizable.

Alternative Derivation of $F^{\mu\nu}$

Since D_μ is gauage invariant, the commutator with itself $[D_\mu, D_\nu]$ is also gauge invariant. So, under the local gauge group, the commutator transforms as

$$g: \ [D_\mu, D_\nu]\psi(x) \to e^{i\alpha(x)}[D_\mu, D_\nu]\psi(x). \tag{266}$$

Plugging in the expression for the covariant derivative and working out the commutator, we get

$$[D_\mu, D_\nu]\psi(x) = [\partial_\mu, \partial_\nu]\psi + ie([\partial_\mu, A_\nu] - [\partial_\nu, A_\mu])\psi - e^2[A_\mu, A_\nu]\psi. \tag{267}$$

The commutators of ∂ and A with themselves are zero. Therefore,

$$[D_\mu, D_\nu]\psi(x) = ie(\partial_\mu A_\nu - \partial_\nu A_\mu)\psi(x) \tag{268}$$

The commutator of the covariant derivative is the *spacetime curvature tensor*.

$$[D_\mu, D_\nu]\psi(x) = ieF_{\mu\nu}(x). \tag{269}$$

Lecture 11: Nonabelian Gauge Theory (Yang-Mills)

As a recap of abelian gauge theory, a *gauge theory* is a theory that is invariant under a group $\mathcal{G}$ of local symmetry transformations, which act independently at each point in spacetime $\mathcal{M}_{1,3}$.

In contrast to local symmetry groups, a global symmetry group that we have dealt with extensively in the Poincaré group, which acts on all of spacetime, making such transformations dependent on spacetime location.

In the context of the Dirac spinor and fermionic field theories, consider the local phase transformation

$$\psi(x) \to e^{i\alpha(x)}\psi(x) \tag{270}$$

Where we assumed that the phase function $\alpha(x)$, $x \in \mathcal{M}_{1,3}$, is differentiable, and maps Minkowski space to the radial unit interval, $\alpha(x) : \mathcal{M}_{1,3} \to [0, 2\pi)$.

The new, larger, more constrained symmetry group that we are building an invariant theory under is Poincaré group plus the local gauge group $\mathcal{G}$. The only terms invaraint under this new symmetry group that we found fit for the Lagrangian density were

$$\bar{\psi}\psi \text{ and } (\bar{\psi}\gamma^\mu\psi)^2. \tag{271}$$

In order compare two independent points of spacetime in this theory, we introduced dynamics in the form of the covariant derivative. The covariant derivative was defined in terms of an auxiliary gauge field A_μ

$$\partial_\mu \to \partial_\mu - ieA_\mu. \tag{272}$$

Call $D_\mu = \partial_\mu - ieA_\mu$, and then we have the additional invariant term $\bar{\psi}\not{D}\psi$ to include in the Lagrangian density under this new representation of the derivative

$$\mathcal{L} = \bar{\psi}\not{D}\psi + F^{\mu\nu}F_{\mu\nu}. \tag{273}$$

Where $F^{\mu\nu}$ is the spacetime curvature tensor that includes derivatives of the gauge field A_μ.

Nonabelian Gauge Theory

Consider the gauge group to be the special unitary group of 2×2 matrices $SU(2)$. Note that for full generality, we should consider arbitrary connected Lie groups, but $SU(2)$ will get us almost all of the story, and the representation theory of $SU(2)$ groups can be used to find the representation theory of general

Lie groups.

What is the local gauge group of $SU(2)$?

This theory must be invariant under group transformations $V(x) \in SU(2)$, where $x \in \mathcal{M}_{1,3}$. An element of this unitary group $V(x)$ has the following form and constraints

$$V(x) = \begin{pmatrix} v_{00}(x) & v_{01}(x) \\ v_{10}(x) & v_{11}(x) \end{pmatrix} \tag{274}$$

$$\begin{aligned} \sum_{j,k=0}^{1} |v_{jk}|^2 &= 2 \\ V^\dagger(x)V(x) &= \mathbb{I} \\ \det(V(x)) &= 1 \end{aligned}$$

We need to choose how $V(x)$ acts on a field. Introduce two independent spinor fields $\psi_0(x)$ and $\psi_1(x)$ that form a new basis, under which the new theory must be invariant,

$$\psi_j(x) = \sum_{k=0}^{1} v_{jk}(x)\psi_k(x). \tag{275}$$

Build the doublet field as an 8×1 vector that Poincaré-transforms like two independent spinors

$$\Psi(x) = \begin{pmatrix} \psi_0(x) \\ \psi_1(x) \end{pmatrix}. \tag{276}$$

How does the local gauge group, with element $g \in \mathcal{G}$, act on the doublet field?

$$g:\ \Psi(x) \to \begin{pmatrix} v_{00}(x)\cdot\mathbb{I}_{4\times4} & v_{01}(x)\cdot\mathbb{I}_{4\times4} \\ v_{10}(x)\cdot\mathbb{I}_{4\times4} & v_{11}(x)\cdot\mathbb{I}_{4\times4} \end{pmatrix} \begin{pmatrix} \psi_0(x) \\ \psi_1(x) \end{pmatrix}. \tag{277}$$

The invariant terms we can construct from the doublet field are

$$\bar{\Psi}\Psi = \sum_{j=0}^{1} \bar{\psi}_j\psi_j \text{ and } (\bar{\Psi}\Gamma^\mu\Psi)^2 \text{ , where } \Gamma^\mu = \begin{pmatrix} \gamma^\mu & 0 \\ 0 & \gamma^\mu \end{pmatrix} \tag{278}$$

As in the abelian case, build the covariant derivative by introducing the *parallel transporter* $U(y,x) \in SU(2)$ and going to a representation of the local gauge group. Under the local gauge transformation

$$U(y,x) \to V(y)U(y,x)V^\dagger(x). \tag{279}$$

The covariant derivative is then defined to be

$$n^\mu D_\mu \Psi \equiv \lim_{\epsilon \to 0} \frac{1}{\epsilon}(\Psi(x+\epsilon n) - U(x+\epsilon n, x)\Psi(x)) \tag{280}$$

Where we note that $U(x+\epsilon n, x)$ is a 2×2 matrix depending on two different spacetime locations. To ensure locality of the theory, we only need to know $U(y,x)$ for $y \simeq x$.

Stepping back, suppose that we have some element of the gauge group $U \in SU(2)$, which we assume is differentiable. Note that this U is not the parallel transporter yet.

Since U is unitary, and recalling that exponentiated unitary elements close to the identity are elements of the underlying Lie algebra, let

$$U = e^{iA} \tag{281}$$

Where A is a Hermitian matrix, such that $A^\dagger = A$ and Tr(A) = 0.

Though it is not necessary, we anticipate computation in the future, and choose a basis. Namely, we choose the 2×2 Pauli spin matrices as a basis, noting that any 2×2 Hermitian traceless matrix can be written as a combination of the three Pauli matrices. In the Pauli basis, the Hermitian matrix is

$$A = \sum_{j=1}^{3} \frac{1}{2} \alpha^j \sigma^j. \tag{282}$$

The Pauli matrices are

$$\sigma^1 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma^2 = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \text{ and } \sigma^3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{283}$$

And obey the Lie bracket

$$[\sigma^j, \sigma^k] = i\epsilon^{jk}{}_l \sigma^l. \tag{284}$$

So, to specify the Hermitian matrix A, we need three real numbers $\alpha^j \in \mathbb{R}$.

Consider some other Hermitian, traceless matrix B, which is to zeroth order equal to the identity, and to first order is proportional to the gauge field. Then the parallel transporter constructed from B has the form

$$U(x+\epsilon n, x) = e^{iB(x;n,\epsilon)} = \mathbb{I}_{2\times 2} + \sum_{j=1}^{3} ig\epsilon n^\mu A^j_\mu \frac{\sigma^j}{2} + \mathcal{O}(\epsilon^3) \tag{285}$$

Then, in the chosen Pauli basis, the covariant derivative is an 8×8 matrix and is defined as

$$n^\mu D_\mu \Psi \equiv \lim_{\epsilon \to 0} \frac{1}{\epsilon}(\Psi(x+\epsilon n) - U(x+\epsilon n, x)\Psi(x)) \tag{286}$$

Where

$$D_\mu = \partial_\mu - igA^j_\mu(x)\frac{\sigma^j}{2}. \tag{287}$$

The coefficient field $A^j_\mu(x)$ is not arbitrary and must obey transformation laws of the local gauge group, as well as give a representation of the local gauge group determined by the action of the local gauge group on the parallel transporter

$$V: U(x+\epsilon n, x) \to V(x+\epsilon n)U(x+\epsilon n, x)V^\dagger(x) \tag{288}$$

$$= V(x+\epsilon n)\left(\mathbb{I} + ig\epsilon n^\mu A^j_\mu \frac{\sigma^j}{2} + \mathcal{O}(\epsilon^3)\right)V^\dagger(x) \tag{289}$$

Calculating the action of $V(x+\epsilon n)$ and $V^\dagger(x)$, the first order term in ϵ becomes (**Exercise**)

$$\text{L.G.}: A^j_\mu(x)\frac{\sigma^j}{2} \to V(x)\left(A^j_\mu \frac{\sigma^j}{2} + \frac{i}{2}\partial_\mu\right)V^\dagger(x). \tag{290}$$

Hint: $V(x+\epsilon n)V^\dagger(x) = [(1+\epsilon n^\mu \partial_\mu)V(x)]\,V^\dagger(x) + \mathcal{O}(\epsilon^2)$.

Next, to compute $V(x)\partial_\mu V^\dagger(x)$, we will take the infinitesimal approach. Recall that for the abelian case, we just had the phase factor $\alpha(x)\partial_\mu\alpha^\dagger(x)$, which was just a number, but now with $V(x)$ we have a 2×2 matrix with an 8×8 representation.

When $V(x)$ is infinitesimally close to the identity, we know that it is some exponential factor in the Pauli sigma matrices

$$V(x) = e^{i\alpha^j(x)\frac{\sigma^j}{2}} \tag{291}$$

Where $\alpha^j(x)$ are small numbers. Applying the Taylor expansion with respect to $\alpha(x)$, the action of $V(x)$ on the partial derivative is

$$V(x)\partial_\mu V^\dagger(x) = (\mathbb{I} + i\alpha^j\frac{\sigma^j}{2})\partial_\mu(\mathbb{I} - i\alpha^j\frac{\sigma^j}{2}) \tag{292}$$

$$= -i\frac{\partial\alpha^j}{\partial x^\mu}\frac{\sigma^j}{2} + \mathcal{O}(\alpha^2) \tag{293}$$

Then, under the local gauge transformation, the gauge field and sigma matrices transform as

$$\text{L.G.}: A^j_\mu(x)\frac{\sigma^j}{2} \to A^j_\mu(x)\frac{\sigma^j}{2} + \frac{1}{g}(\partial_\mu\alpha^j(x))\frac{\sigma^j}{2} + i\left[\alpha^j(x)\frac{\sigma^j}{2}, A^k_\mu(x)\frac{\sigma^k}{2}\right]. \tag{294}$$

Now we can see the infinitesimal local gauge transformation does to the covariant derivative of the doublet spinor field $\Psi(x)$

$$\text{L.G.} : \Psi(x) \to \left(\mathbb{I} + i\alpha^j(x)\frac{\sigma^j}{2} \right) \Psi(x) \tag{295}$$

$$\text{L.G.} : D_\mu\Psi(x) \to \left(\partial_\mu - igA^j_\mu(x)\frac{\sigma^j}{2} - i(\partial_\mu\alpha^j(x))\frac{\sigma^j}{2} + g\left[\alpha^j(x)\frac{\sigma^j}{2}, A^k_\mu(x)\frac{\sigma^k}{2}\right] \right) \left(1 + i\alpha^j(x)\frac{\sigma^j}{2} \right) \Psi(x) \tag{296}$$

To first order in α, the right-hand side of the infinitesimal transformation becomes

$$\text{L.G.} : D_\mu\Psi(x) \to \left(1 + i\alpha^j(x)\frac{\sigma^j}{2} \right) D_\mu\Psi(x) \tag{297}$$

$$= V(x)D_\mu\Psi(x) \tag{298}$$

Where, for the physicist, ignoring issues of connectivity with the local gauge group (will need *gauge fixing*), we make the "big" gauge transformation (last line) by exponentiating (e.g., $(1 + \frac{x}{n})^n = e^x$).

Now we need to build a Langrangian density term that gives dynamics to the gauge field $A^j_\mu(x)$. Recall the commutator $[D_\mu, D_\nu]$, the curvature of the $SU(2)$ fibre bundle, which is local gauge invariant, and involves derivatives of A^j_μ. This transforms under local gauge as

$$\text{L.G.} : [D_\mu, D_\nu]\Psi(x) \to V(x)[D_\mu, D_\nu]\Psi(x). \tag{299}$$

In the nonabelian case, the commutator has the form, using the fact that mixed partial derivatives commute (**Exercise**)

$$[D_\mu, D_\nu] = igF_{\mu\nu}{}^j\frac{\sigma^j}{2} \tag{300}$$

$$= ig\left(\partial_\mu A^j_\nu\frac{\sigma^j}{2} - \partial_\nu A^j_\mu\frac{\sigma^j}{2} - ig\left[A^j_\mu\frac{\sigma^j}{2}, A^k_\nu\frac{\sigma^k}{2}\right] \right) \tag{301}$$

So, in the nonabelian case, the curvature tensor $F_{\mu\nu}$ depends quadratically on A^j_μ, whereas in the abelian case it was linearly dependent. Therefore, the invariant term $\sim F^{\mu\nu}F_{\mu\nu}$ yields cubic and quartic terms in the Lagrangian density, making an "interacting" theory.

Under the local gauge, the curvature tensor transforms as similarity

$$\text{L.G.} : F_{\mu\nu}{}^j\frac{\sigma^j}{2} \to V(x)(F_{\mu\nu}{}^j)V^\dagger(x). \tag{302}$$

By the similarity transformation, we can build an invariant Lorentz scalar with the trace of the invariant term (**Exercise**, and note that $\sigma^j \sigma^k$ contributes δ_{jk})

$$\mathrm{Tr}\left(\left(F_{\mu\nu}{}^{j}\frac{\sigma^j}{2}\right)\left(F_{\mu\nu}{}^{k}\frac{\sigma^k}{2}\right)\right) = \frac{1}{8}(F_{\mu\nu}{}^{j})^2. \tag{303}$$

The (classical) Langrangian density for the nonabelian gauge theory, invariant under local gauge and Poincaré transformations is

$$\mathcal{L} = \bar{\Psi}(i\not{D} - m)\Psi - \frac{1}{4}(F_{\mu\nu}{}^{j})^2. \tag{304}$$

Note that in the abelian case (theory of QED), the dynamics in the Lagrangian density of the gauge field were quadratic. Considering only the gauge field term, with no matter (fermions), essentially results in the wave equation.

$$F_{\mu\nu}^2 = (\partial_\mu A_\nu - \partial_n u A_\mu)^2 \tag{305}$$

In the nonabelian case, they are cubic and quartic, since there is the commutator of the gauge fields.

The Lagrangian density can be quantized in two ways:

1. Perturbatively via path integrals
2. Non-perturbatively via lattice discretization.

Lecture 12: Quantization of Gauge Theories

Peskin and Schroeder, page 294

Recall the classical Lagrangian density $\mathcal{L} = \bar{\psi}(i\not{D} - m)\psi - \frac{1}{4}F_{\mu\nu}{}^{a}F^{\mu\nu,\,a}$, where a denotes individual fields, that we crafted to be invariant under the local gauge symmetry group $SU(2)$. We introduced the "helper" gauge field A^a_μ which manifests in the terms of the spacetime curvature tensor , or the curvature of the $SU(2)$ fibre bundle, $F_{\mu\nu} = -i[D_\mu, D_\nu]$, where $D_\mu = \partial_\mu - igA^a_\mu \frac{\sigma^a}{2}$ is the covariant derivative.

This Langrangian density represents a nontrivial dynamical system that is invariant under the local gauge symmetry group and endows fermions, as well as other fields, with dynamics. Unlike many other effective classical theories, this one is not quadratic in its fields and yields nonlinear equations of motion (e.g., instanton and soliton solutions).

We now quantize this gauge theoryto build a quantum theory that is invariant under the Poincaré and local gauge symmetry groups by finding the correct representation that has this Lagrangian density as its effective classical limit.

Two problems that arise for the gauge theories are (1) the classical theory ($\mathcal{L}_0$) is already nonlinear, and (2) there are lots of symmetries, global and local. Local symmetries are represented by copies of $SU(2)$ acting independently of each other at each spacetime location.

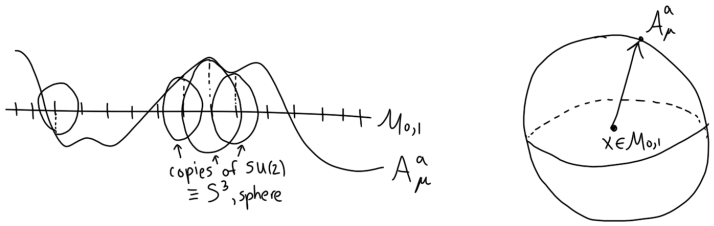
Two approaches of quantization that we will explore include (1) an analytic, but naive *path integral quantization* $\int \mathcal{D}A\mathcal{D}\psi\mathcal{D}\bar{\psi}e^{iS}$ that is good for high-energy scattering processes, but not for calculating ground state correlation functions. The mathematical rigor of this approach is a current topic of research. And (2) a computational route of *lattice quantization*, which makes dealing with nonlinearity "easy", but it loses Poincaré invariance in the process. Note that there is also the route of canonical quantization, but we will not bother with that here.

Path Integral Quantization of Gauge Theories

First, some words on the space of all gauge fields A^a_μ: note that $A^a_\mu \frac{\sigma^a}{2}$ and $A^a_\mu \frac{\sigma^a}{2} + \frac{1}{g}(\partial_\mu \alpha^a)\frac{\sigma^a}{2} + i[\alpha^b \frac{\sigma^b}{2}, \alpha^c \frac{\sigma^c}{2}]$ are *gauge equivalent*, meaning that there exists an infinite number of A^a_μ's with the same path integrand e^{iS}, causing the path integral to result in infinity, since we are redundantly integrating over a continuous infinity of physically equivalent field configurations.

Consider the gauge field A^a_μ, a list of 12 numbers in (3+1)-dimensional spacetime, on $(0+1)$-dimensional spacetime $\mathcal{M}_{0+1}$ (e.g., a line), where A^a_μ is still a list

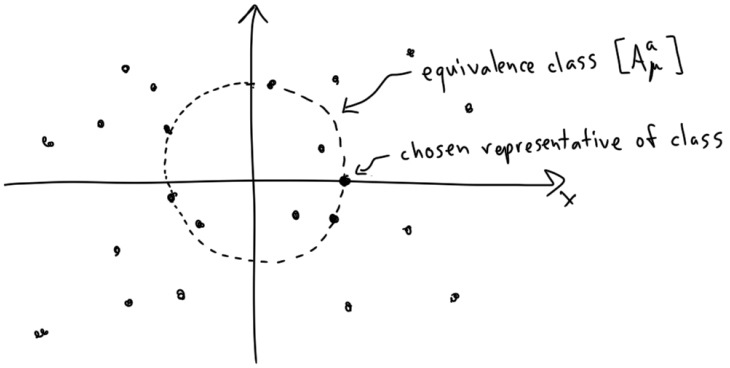
of three numbers. Now, $SU(2)$ acts independently on A^a_μ at each point in $\mathcal{M}_{0+1}$. This action is tantamount to multiplying by a phase on a sphere S^3 at each spacetime location, since $SU(2)$ is parameterized by four numbers, the coefficients of the quaternions with norm equal to one. Below is a schematic of the space of all A^a_μ's.



Schematic of gauge field configurations, independent copies of $SU(2)$ acting on $(0+1)$-dimensional spacetime.

So A^a_μ is like a vector on S^3 at each spacetime location and is a possible configuration in $\mathcal{M}_{0+1}$. These configurations form an equivalence class $[A^a_\mu]$ defined by the local gauge group $\mathcal{G}$. Addressing the problem of very many symmetries, global and local, in the path integral approach, there are very many equivalence classes, or configurations, to sum over, but we'd like to just choose one configuration.

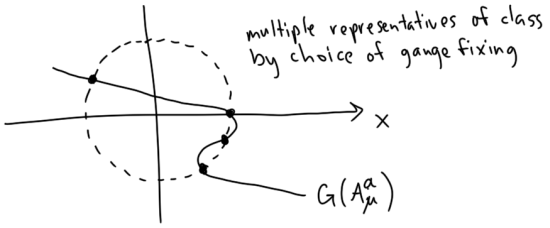
Another, more common schematic to demonstrate the action of gauge groups is to consider rotations in $SO(2)$, where our theory is a zero-dimensional theory invariant under $SO(2)$. Configurations in this schematic are just points in a two-dimensional space with equivalence classes defined by circles centered at the origin. See schematic below.



Schematic of gauge field configurations and equivalence classes represented by rotations in $SO(2)$.

In choosing a representative point per equivalence class, we enforce the *gauge fixing condition*. A good choice of representative may be the point that crosses the horizontal axis. We then integrate over the space of the chosen representatives. This effectively reduces the size of the configuration space and makes the path integral much more tractable.

Note that we only pick one representative per equivalence class, but there can be more than one depending on the choice of gauge fixing condition, where the gauge fixing condition if a function that crosses the equivalence class circle more than once. This is called the *Gribov ambiguity*, and commonly happens in the Coulomb gauge.



Example schematic of Gribov ambiguity, where more than one representative of the equialvence class is chosen by the guage fixing condition.

Gauge Fixing Condition

Recall that the path integral approach is made difficult by the fact that we are redundantly integrating over a continuous infinity of physically equivalent field

configurations. By applying a gauge fixing condition, we isolate the intersting part of the integral and count each distinct physical configuration only once. Finding the right gauge fixing function G allows us to separate out this overcounting in the path integral and throw it away. Note that we are free to "throw it away" since we are *guessing* a quantum theory.

Choosing the gauge fixing function to be $G(A^a_\mu) = \partial_\mu A^a_\mu - \omega^a$, where ω^a is any scalar field, is a generalization of the Lorentz gauge, and setting this equal to zero is a generalization of the Lorentz gauge condition

$$G(A^a_\mu) = \partial_\mu A^a_\mu - \omega^a = 0. \tag{306}$$

With the gauge fixing function, the way we separate the overcounting is by inserting unity

$$1 = \int \mathcal{D}\alpha \; \delta(G(A^\alpha)) \det\left(\frac{\delta G(A^\alpha)}{\delta \alpha}\right). \tag{307}$$

Breaking equation this down, we are performing a path integral over all possible gauge transformations, and picking out only the $G(A^\alpha)$ that equals zero, obeying the gauge fixing condition, choosing a single representative of the equivalence class. The determinant is called the *Faddeev-Popov determinant*. The notation A^α indicates the locally gauge transformed gauge field

$$(A^\alpha)^a_\mu = A^a_\mu + \frac{1}{2}\partial_\mu \alpha^a + f^{abc} A^b_\mu \alpha^c \tag{308}$$

$$= A^a_\mu + \frac{1}{g} D_\mu \alpha^a \tag{309}$$

Where f^{abc} are the structure constants from the Pauli spin matrix commutation relations

$$\left[\frac{\sigma^a}{2}, \frac{\sigma^b}{2}\right] = i f^{abc} \frac{\sigma^c}{2}. \tag{310}$$

This way of writing 1 is the continuous, functional generalization of that for discrete, many-variable n-dimensional vectors

$$1 = \left(\prod_{j=1}^{n} \int da_j\right) \delta^{(n)}(\mathbf{g}(\mathbf{a})) \det\left(\frac{\partial g_j}{\partial a_k}\right) \tag{311}$$

Where the determinant here is the Jacobian of the change of variables. By change of variables, this can be written as (**Exercise**)

$$1 = \left(\prod_{j=1}^{n} \int db_j\right) \delta^{(n)}(\mathbf{b}). \tag{312}$$

Inserting the contiuous, functional version of one into the path integral, we now have an expression that integrates over the equivalence classes but "sucks out" the overcounting to just one representative of the class

$$\int \mathcal{D}\alpha \int \mathcal{D}A\mathcal{D}\psi\mathcal{D}\bar{\psi}\ \delta(G(A^\alpha))\det\left(\frac{\delta G(A^\alpha)}{\delta\alpha}\right)e^{iS}. \tag{313}$$

Evaluating the Faddeev-Popov determinant in our choice of the Lorentz gauge (**Exercise**)

$$\det\left(\frac{\delta G(A^\alpha)}{\delta\alpha}\right) = \det\left(\frac{1}{g}\partial^\mu D_\mu\right) \tag{314}$$

$$= \int \mathcal{D}c\mathcal{D}\bar{c}\ e^{i\frac{1}{g}\int d^4x\,\bar{c}(-\partial^\mu D_\mu)c} \tag{315}$$

Where, recall from the study of fermions, we have used auxiliary Grassman-valued, scalar, spin-0 fields c and $\bar{c}$. These fields are non-physical and must disappear form the final results: *Faddeev-Popov ghosts* or *ghosts*.

Inserting the ghost expression for the determinant into the path integral, we now have

$$\int \mathcal{D}\alpha \int \mathcal{D}A \int \mathcal{D}c\mathcal{D}\bar{c} \int \mathcal{D}\psi\mathcal{D}\bar{\psi}\ \delta(\partial^\mu A^a_\mu - \omega^a)e^{i(S+\frac{1}{g}\int d^4x\,\bar{c}(-\partial^\mu D_\mu)c)}. \tag{316}$$

Integrate out the delta functional, since ω^a is arbitrary, using an Gaussian integral over ω^a with coefficient $\xi \in [0,1]$, and calling $S' = S + \frac{1}{g}\int d^4x\,\bar{c}(-\partial^\mu D_\mu)c$,

$$\int \mathcal{D}\omega\, e^{-i\int d^4x\,\frac{1}{2}\xi(\omega^a)^2}\left(\int \mathcal{D}...\right)\delta(\partial^\mu A^a_\mu - \omega^a)e^{iS'} \tag{317}$$

Since ω^a is an arbitrary, gauge-fixing, scalar function that takes one representative of each independent equivalence class, the full path integral will now be independent of ω^a, and we have the form

$$N(\xi)\int \mathcal{D}\alpha \int \mathcal{D}A \int \mathcal{D}c\mathcal{D}\bar{c} \int \mathcal{D}\psi\mathcal{D}\bar{\psi}\ e^{i\int d^4x\,\mathcal{L}'}. \tag{318}$$

The Lagrangian $\mathcal{L}'$ has the form

$$\mathcal{L}' = \bar{\psi}(i\not{D} - m)\psi - \frac{1}{4}(F^a_{\mu\nu})^2 + \frac{1}{2}\xi(\partial^\mu A^a_\mu)^2 + \frac{1}{g}\bar{c}^a(-\partial^\mu D^{ab}_\mu)c^b. \tag{319}$$

Note that the integral over α blows up to infinity, but in correlation functions we always have ratios of the path integrals and the $N(\xi)\cdot\infty$'s will cancel out.

Remaining questions include

- Does the path integral above even define a quantum theory, and is it invariant under Poincaré and local gauge symmetry group transformations?

  See the work of 't Hooft.

- Do the ghosts c and $\bar{c}$ vanish from the processes?

  Feynman diagrams will clear up this concern.

- Is the Lagrangian density (theory) $\mathcal{L}'$ renormalizable?

  Also see the work of 't Hooft.

Lecture 13: Quantization of Nonabelian Gauge Theory

In the path integral quantization of gauge theories, we started by guessing a quantum theory by integrating the action functional over fermion fields ψ, $\bar{\psi}$ and gauge boson fields A to calculate transition amplitudes.

We noticed that there are many gauge-equivalent configurations of the gauge field that causes naive divergences, due to overcounting, in the Feynman diagram expansion. To tame this infinity, we fixed the gauge by choosing a single representative from each gauge equivalence class by inserting unity in terms of the delta functional. This introduced a nonlinear Jacobian term $\det\left(\frac{\delta G(A^\alpha)}{\delta\alpha}\right)$, where A^α denotes the gauge-fixed gauge field.

To compute this determinant, we introduced the auxiliary (classical, scalar, spinless) Grassman-valued fields $c(x)$, called Faddeev-Popov ghosts. These fields are unphysical, and must drop out during calculation. Note that the PCT-thereom does not apply to ghosts. Then transition amplitude is then defined as

$$\langle \Phi_f | \, U \, | \Phi_i \rangle \equiv \int \mathcal{D}A\mathcal{D}\psi\mathcal{D}\bar{\psi}\mathcal{D}c\mathcal{D}\bar{c} \; e^{iS}. \tag{320}$$

Where the ghosts obey the anticommutation relation as Grassman-valued fields $\{c(x), c(y)\} = 0$, and the Lagrangian density is

$$\mathcal{L} = -\frac{1}{4}(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu)^2 + \frac{1}{2}\xi(\partial^\mu A^a_\mu)^2 + \bar{\psi}(i\not{\partial} - m)\psi + \bar{c}^a(-\partial^\mu\partial_\mu)c^a + \mathcal{L}_{int}(g). \tag{321}$$

Note that the interacting bits, including the bits from the covariant derivative D_μ are absorbed into $\mathcal{L}_{int}(g)$, and everything else in the expression above is the free theory with partial derivatives ∂_μ.

We now push forward under the belief that the perturbatively defined theory above is representative of a quantum theory, and expand in powers of the interaction term g to define processes and Feynman rules of the **Yang-Mills theory**.

Propagators (free theory, $g = 0$ part)

The fermion propagator contributes

$$\langle \bar{\psi}_{j\alpha}(x)\hat{\bar{\psi}}_{l\beta}(y)\rangle = \int \frac{d^4k}{(2\pi)^4}\left(\frac{i}{\not{k} - m}\right)_{\alpha\beta}\delta_{jl}e^{-ik\cdot(x-y)}. \tag{322}$$

The gauge boson propagator contributes

$$\langle \hat{A}^a_\mu(x)\hat{A}^b_\nu(y)\rangle = \int \frac{d^4k}{(2\pi)^4}\ \left(\eta_{\mu\nu} - (1-\xi)\frac{k_\mu k_\nu}{k^2}\right)\delta_{ab}\frac{e^{-ik\cdot(x-y)}}{k^2+i\epsilon}. \tag{323}$$

The ghost propagator contributes

$$\langle \hat{c}^a(x)\hat{\bar{c}}^b(y)\rangle = \int \frac{d^4k}{(2\pi)^4}\ \frac{i}{k^2}\delta_{ab}e^{-ik\cdot(x-y)}. \tag{324}$$

Recall that a, $b = 1, 2, 3$ and μ, $\nu = 0, 1, 2, 3$.

Vertices (interacting theory, $g \neq 0$ part)

The interaction of two fermions and one gauge boson contributes

$$ig\gamma^\mu\frac{\sigma^a}{2}. \tag{325}$$

The interaction of three gauge bosons contributes

$$gf^{abc}\left(\eta^{\mu\nu}(k-p)^\rho + \eta^{\mu\rho}(p-q)^\mu + \eta^{\rho\mu}(q-k)^\nu\right). \tag{326}$$

The interaction of four gauge bosons contributes

$$-ig(f^{abc}f^{cde}(\eta^{\mu\rho}\eta^{\nu\sigma}-\eta^{\mu\sigma}\eta^{\nu\rho})+f^{ace}f^{bde}(\eta^{\mu\nu}\eta^{\rho\sigma}-\eta^{\mu\sigma}\eta^{\nu\rho})+f^{ade}f^{bce}(\eta^{\mu\nu}\eta^{\rho\sigma}-\eta^{\mu\rho}\eta^{\nu\sigma})). \tag{327}$$

The interaction of one gauge boson with two ghosts contributes

$$gf^{abc}p^\mu. \tag{328}$$

(**Exercise**) And the rest of the Feynman rules for Yang-Mills theory: symmetry factors, signs, and conservations laws.

Big Question 1

Can we interpret this as a quantum theory? In other words, does this theory implicitly define the Hermitian operators $\hat{A}^\alpha$, $\hat{\psi}$, and $\hat{c}$?

Yes, it is a quantum theory, subject to a cutoff.

In more detail, what are the possible modes of failure for the theory to not be a valid quantum theory?

The first mode of failure that can occur is the time evolution operator not being unitary, or time translation not being a unitary process. To confirm that time translation in this theory is a unitary process, check that the correlation function is symmetric under the group of Poincaré transformation (**Exercise**). This also

implies that probability is conserved.

The second mode of failure occurs if, after building the Fock space, we get negative norm states, such that the inner product of the system eigenstates is less than zero $\langle\Psi|\Psi\rangle < 0$. If this happens, then the Hamiltonian is not positive definite and we do not have a proper Hilbert space.

Note that there is a type of exception to this rule, which is a topic of *current research*. Negative states may be able to be modded out by their subspaces to produce an effective quantum theory from the remaining subspaces with positive inner products. These remaining states are the physical, operationally defined states, and they form a convex cone from an operator algebra. The inner product in this state subspace forms a "Hilbert subspace", where $\langle\Psi_{\text{phys}}|\Psi_{\text{phys}}\rangle > 0$.

Big Question 2

Is this theory, as a quantum theory, renormalizable?

Yes, thanks to the work of t'Hooft and Veltman in the REGULARIZATION AND RENORMALIZATION OF GAUGE FIELDS. They found a smart cutoff to impose that renders infinite integrals finite and retains Lorentz and gauge invariance through the technique called *dimensional regularization*.

Note that different cutoffs can reveal different hypotheses about reality, and it is generally believed that physics is Lorentz and Poincaré invariant. So, a good cutoff should retain these invariances while also taming infinities of the Feynman diagram integral contributions.

Dimensional Regularization: An Example

Peskin and Schroeder, page 249.

In ϕ^4 theory, as well as others, we encounter loop integrals that diverge because we are imposing too strong of a hypothesis on how the degrees of freedom of the theory behave. For example take the following loop integral that shows up in the Feynman expansion

$$I_n(m) = \int \frac{d^d k}{(2\pi)^d} \frac{1}{(k^2 - m^2 + i\epsilon)^n} \tag{329}$$

Where $n \in \mathbb{Z}$ is the number of undetermined momenta. This produces a naive divergence of the form Λ^d/Λ^{2n}.

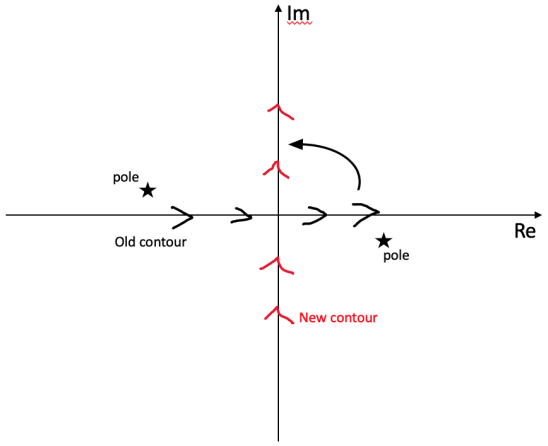
Note that in lower dimensions, these divergences are easier to handle. For example, if $d = 2$, then $n = 1$ tames the infinity. If $d = 3$, n must be greater than

one to tame infinities of the cutoff in just three dimensions. In four dimensions, such as our usual spacetime, $n \geq 2$ is required to make the proper cancellations (e.g., need more undetermined momenta).

Now to evaluate that integral, we separately write out the time component integral and note that there are two poles at $k_0 = \pm(\sqrt{k^2 + m^2} - i\epsilon)$, where k is just the spatial components of the k-vector (e.g., three spatial components of spacetime), and the $i\epsilon$ is Taylor expanded out of the square root,

$$I_n(m) = \int \frac{d^{d-1}k}{(2\pi)^d} \left(\int_{-\infty}^{\infty} dk_0 \, \frac{1}{(k_0^2 - k^2 - m^2 + i\epsilon)^n} \right). \tag{330}$$

Now, rotate the contour from running along the real axis $\Re(k_0)$ to the imaginary axis $\Im(k_0)$, effectively changing variables $k_0 \to ik_0$. Avoiding the poles, everything in the region is analytic, meromorphic and we can do this rotation.



Contour rotation from real-axis of k_0 to the imaginary axis of k_0.

Our integral is now

$$I_n(m) = -i \int \frac{d^{d-1}k}{(2\pi)^d} \int_{-\infty}^{\infty} dk_0 \, \frac{1}{(k_0^2 + \omega^2)^n}. \tag{331}$$

Where $\omega^2 = k^2 + m^2$. Reabsorb the time-component integral into the rest of the integrals, which we can interpret as d Euclidean integrals, with the Euclidean metric, that will now exhibit spherical symmetry. Enter spherical polar coordinates

$$I_n(m) = -i \int_{\text{Euclidean}} \frac{d^d k}{(2\pi)^d} \frac{1}{(k^2+m^2)^n} \tag{332}$$

$$= -\frac{i}{(2\pi)^d} \int d\Omega_d \int_0^\infty dr \, \frac{r^{d-1}}{(r^2+m^2)^n} \tag{333}$$

$$I_n(m) = -\frac{i}{(2\pi)^d} \frac{(2\pi)^{\frac{d}{2}}}{\Gamma(\frac{d}{2})} \int_0^\infty dr \, \frac{r^{d-1}}{(r^2+m^2)^n}. \tag{334}$$

Make the change of variables $x = \frac{m^2}{r^2+m^2}$

$$I_n(m) = \frac{m^{d-2n}}{(2\pi)^{\frac{d}{2}}\Gamma(\frac{d}{2})} \int_0^1 dx \, x^{n-\frac{d}{2}-1}(1-x)^{\frac{d}{2}-1}. \tag{335}$$

This integral is a *Beta function*!

$$I_n(m) = \frac{m^{d-2n}}{(2\pi)^{\frac{d}{2}}\Gamma(\frac{d}{2})} \frac{\Gamma(n-\frac{d}{2})\Gamma(\frac{d}{2})}{\Gamma(n)} = \frac{m^{d-2n}}{(2\pi)^{\frac{d}{2}}} \frac{\Gamma(n-\frac{d}{2})}{\Gamma(n)} \tag{336}$$

The gamma function has poles at negative integers, and this has divergences at $n = 0, -1, -2, \ldots$. The Feynman diagram expansion gives us n, and we are stuck with it.

What if d is not a integer? This is the trick of dimensional regularization, as this renders the divergences finite.

Let $d = 4 - \epsilon$, $\epsilon > 0$, and study the behavior of the integral solutions as $\epsilon \to 0$. As an example, consider the $n - 2$ case (**Exercise**)

$$I_2(m) = \frac{m^{-\epsilon}}{(2\pi)^{2-\frac{\epsilon}{2}}} \Gamma\left(\frac{\epsilon}{2}\right) \tag{337}$$

$$= \frac{1 - \epsilon \log(m)}{4\pi^2(1 - \frac{\epsilon}{2}\log(2\pi))} \left(\frac{2}{\epsilon} - \gamma\right) + \mathcal{O}(\epsilon^2) \tag{338}$$

Where γ is the Euler-Mascheroni constant. This diverges as $\epsilon \to 0$, as expected, with the "bad bit" of $\frac{2}{\epsilon}$.

If we compare this with an integral that we've seen before, we can seee what kind of cutoff the free parameter ϵ is. Recall the diagram where we applied the momentum cutoff $|k| < \Lambda$. So, $\frac{1}{\epsilon} \sim \Lambda$.

Lecture 14: Quantization of Nonabelian Gauge Theory, Cont.

Working with nonabelian gauge theories, we've written down a Lagrangian density for Yang-mills theory, and after the gauge-fixing procedure of the Lagrangian density via path integrals and the tricks of Faddeev and Popov, we are ready to do some calculations.

The first topic to discuss is the *beta function* or *renormalization group equation*, which tells us how theories behave at low and high energies. The second topic to discuss is the departure from path integrals and gauge fixing to methods of lattice regulators, or cutoffs, for doing calculations in nonabelian gauge theory.

Renormalization of Nonabelian Gauge Theories

We begin by reviewing renormalizability of quantum theories and the renormalization group equation, which tells us how coupling constants depend on the cutoff and how to adjust the coupling constants to match low and high energy predictions.

Recall that a quantum theory $\hat{H}(z_1, \ldots, z_n; \Lambda)$, where Λ is the cutoff, data defined in a list of all the degrees of freedom of the theory, is renormalizable is it leads to finite predictions for all operationally well-defined observables. The expectation values of these observables must produce the same predictions for different choices of the chosen cutoff

$$\langle \hat{A}_j \rangle (z_1, \ldots, z_n; K_c) \equiv f_j(z_1, \ldots, z_n; K_c) = \alpha_j^{\text{obs.}} \tag{339}$$

Where K_c is a particular choice of cutoff: the length of list Λ, for example. This relationship of expectation values with different cutoffs yielding the same observed quantities is achieved in a renormalizable theory by allowing the coupling constants to depend on the cutoff $z_i = z_i(K_c)$.

Apply the above equation to the Green's function, n-point correlation functions, where the dependency of the coupling constants and cutoff are implicit to the vacuum state $|\Omega\rangle$ and the dynamics of the field operators

$$G^{(n)}(x_1, \ldots, x_m; K_c) \equiv \langle \Omega | \, \mathcal{T}[\hat{\phi}(x_1) \ldots \hat{\phi}(x_m)] \, | \Omega \rangle \tag{340}$$

$$, \text{ where } |\Omega\rangle = |\Omega(z_1, \ldots, z_n; K_c)\rangle \tag{341}$$

$$, \text{ and } \hat{\phi}(x) \equiv e^{-i\hat{H}(z_1,\ldots,z_n;K_c)t} \hat{\phi}(0, \underline{x}) e^{i\hat{H}(z_1,\ldots,z_n;K_c)t}. \tag{342}$$

So, we can equivalently say that a quantum theory $\hat{H}(z_1, \ldots, z_n; \Lambda)$ is renormalizable if the correlation functions produce finite predictions for time-ordered quantities.

Now, to compute the relationship of $G^{(n)}$ to the coupling constants and the cutoff, consider changing the (usually) continuous parameter K_c by an infinitesimal amount. Note that a common choice of cutoff could be $K_c = |p_{\max}|$.

$$dG^{(n)} = \frac{\partial G^{(n)}}{\partial K_c}\delta K_c + \frac{\partial G^{(n)}}{\partial z_j}\delta z_j. \tag{343}$$

So, the coupling constants $z_j(K_c)$ are chosen to fix $G^{(n)}$ with respect to transformations of the cutoff of the form $K_c \to K_c + \delta K_c$, but $G^{(n)}$ should not depend on K_c, and the above expression is equal to zero

To derive the *renormalization group equation* or the *beta function*, set the differential $dG^{(n)}$ to zero, divide by δK_c, and multiply by K_c

$$\left[K_c\frac{\partial}{\partial K_c} + K_c\frac{dz_j}{dK_c}\frac{\partial}{\partial z_j}\right]G^{(n)} = 0 \tag{344}$$

$$\left[K_c\frac{\partial}{\partial K_c} + \beta(z_j)\frac{\partial}{\partial z_j}\right]G^{(n)} = 0. \tag{345}$$

This is the infinitesimal form of the statement for a renormalizable theory that the Green's function shouldn't depend on the cutoff as it is changed, where

$$\beta(z_j) = K_c\frac{dz_j}{dK_c} = \frac{dz_j}{d\ln K_c} \tag{346}$$

Is the renormalization group or beta function, which allows us to compute the coupling constants in terms of the cutoff $z_j = z_j(K_c)$. The behavior of $\beta(z_j(K_c))$ is often used to describe the dependence of z_j on K_c.

Aside: Massless Theories

When we fix the Green's function to be equal to observable quantities, for any choice of cutoff, in a massive theory, we usually demand that the 2-point correlation function $\langle\Omega|\,\hat{\phi}(p)\hat{\phi}(-p)\,|\Omega\rangle$ has a pole at the physical mass of the particle m_{phys}.

Complications arise in massless theories, as this leads to divergences, in the sense that we will end up with expressions like $\infty = \infty$. For massless theories, we instead insist that the 2-point correlation function has a pole at a negative, spacelike momenta $p^2 = -K_c^2 \equiv M^2$ with residue equal to one. This leads to finite predictions when we let $K_c \to \infty$.

Examples of Beta Functions

(1) In ϕ^4 theory, there is one coupling constant λ, the interaction strength, and the beta function is

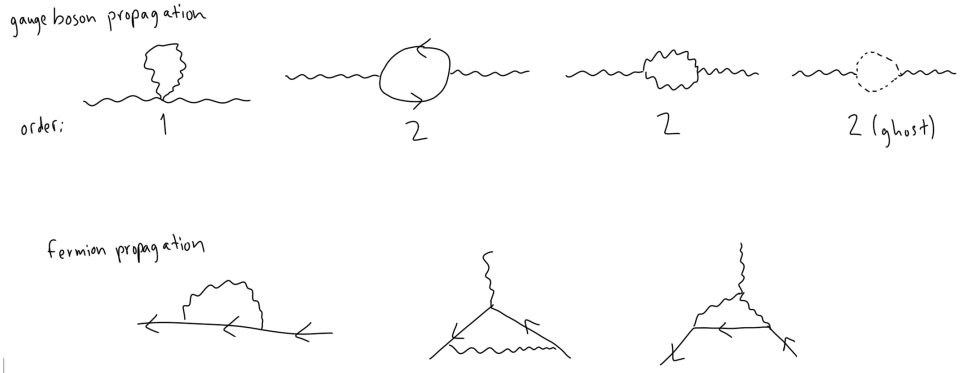
$$\beta(\lambda) = \frac{3\lambda^2}{16\pi^2} + \mathcal{O}(\lambda^3). \tag{347}$$

(2) In quantum electrodynamics (QED), which is an abelian gauge theory, to first loop order, the beta function is

$$\beta(e) = \frac{e^3}{12\pi^2} + \mathcal{O}(e^4)t \tag{348}$$

Where e is the electric charge of the particle.

(3) In Yang-Mills theory, a nonabelian gauge theory, there is more work to do in analyzing the divergences in the propagation of gauge bosons, as well as the propagation of fermions (quarks). The path integrals derived from applying the Yang-Mills theory Feynman rules produce divergences from the following terms in the Feynman expansion. See *Peskin and Schroeder*, section 16.5 *One-Loop Divergences of Non-Abelian Gauge Theory*, pages 521-544, for reference.



To tame the infinities that arise from these diagrams, apply dimensional regularization to maintain Lorentz invariance and insist that each diagram leads to finite quantities. This then tells us what kinds of counter terms we need to add to the Lagrangian density for our theory, and, in turn, how the coupling constants change with respect to the cutoff.

This leads to the beta function for the $SU(N)$ local gauge group

$$\beta(g) = -\frac{g^3}{16\pi^2}\left(\frac{11N}{3} - \frac{2n_f}{3}\right) \tag{349}$$

Where n_f is the number of fermion families. Note that if n_f is small, $\beta(g)$ becomes negative, and that $K_c \to \infty$ as g gets smaller, meaning that our theory approaches being a free theory as $g \to 0$, and we can use perturbation theory for

high-energy processes. This is called *asymptotic freedom* of nonabelian gauge theories.

Low-Energy Physics of Nonabelian Gauge Theories

Results here are thanks to Wilson's *Confinement of Quarks* (1974).

To quantize an abelian or nonabelian gauge theory onto a discrete lattice, Wilson proposed the use of a lattice regulator. This is very challenging, since the gauge group acts on gauge fields like

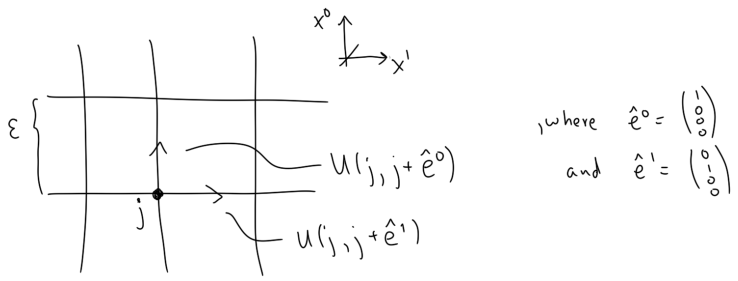
$$A_\mu^j \frac{\sigma^j}{2} \to A_\mu^j \frac{\sigma^j}{2} + \frac{1}{g}(\partial_\mu \alpha^j)\frac{\sigma^j}{2} + i\left[\alpha^k \frac{\sigma^k}{2}, \alpha^l \frac{\sigma^l}{2}\right] \tag{350}$$

And there any many complications in discretizing the derivative $\partial_\mu \alpha^j$ to the lattice.

Wilson recognized that the parallel transporter is the object that allowed us to do derivatives in the first place, and treated the parallel transporter

$$U(j, j + \hat{e}^\mu) \in SU(2) \tag{351}$$

as the fundamental degrees of freedom of the nonabelian gauge theory discretized to the lattice $\epsilon\mathbb{Z}^4$, instead of the gauge field.



Bigger parallel transporters are built via multiplication. For example,

$$U(j, j + 2\hat{e}^\mu) = U(j + \hat{e}^\mu, j + 2\hat{e}^\mu)U(j, j + \hat{e}^\mu) \tag{352}$$

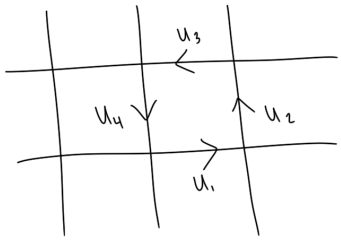
These parallel transporters are 2×2 unitary matrices which populate the list of degrees of freedom, one for each link, or edge, in the classical lattice.

Side: Wilson ran calculations for the dynamics of a gauge field on a $4 \times 4 \times 4 \times 4$ lattice, which requires 4^4 lattice sites $\times 3$ spatial coordinates per lattice site = 768 single-precision floating point numbers per unit time, requiring just 4 kilobytes of RAM. Note that a gigabyte ($\simeq 2^{30}$ bytes) of RAM is capable of storing $\sim 250,000,000$ single-precision floating point numbers, corresponding to almost $100 \times 100 \times 100 \times 100$ lattice.

To quantize the nonabelian gauge theory, Wilson proposed a way to build an action that summed over the plaquettes of the lattice, called *Wilson loops*. In terms of the parallel transporters, the action has the form

$$S[U] = \sum_{\square} \text{Wilson loops} = \sum_{\square} \text{tr}(U_1 \cdot U_2 \cdot U_3 \cdot U_4) \tag{353}$$

Where travelling around a plaquette may look like



With this action, build the path integral $\int \mathcal{D}U\, e^{-iS[U]}$, and Wick rotate into the path integral $\int \mathcal{D}U\, e^{-S[U]}$ to work in imaginary time. Lastly, Monte Carlo sample the path integral.

This approach is actually the best way to get nonperturbative results in quantum field theory, but has its downsides:

Downside (1) is calculating processes in imaginary time is just like doing statistical mechanics with gauge theories at some defined temperature, and this is not good for time-ordered processes (e.g., scattering). Downside (2) is that this is not a quantum theory, as the Wilson loop is a classical configuration.

Hamiltonian Lattice Gauge Theory

An alternative approach to quantization was introduced by Kogut and Susskind in 1975. They proposed a lattice quantum gauge theory, a quantum theory with a Hamiltonian and Hilbert space in which the degrees of freedom live on a lattice, which they argued yields Yang-Mills theory as the lattice spacing goes to zero. This is not proven, but if you can prove it, as well as that the low-energy limit

has a mass cap, you can get a cool $1M!

Recall that classically, each link, or edge, of the lattice is associated to a 2×2 unitary matrix $U \in SU(2)$. In this lattice quantum gauge theory, each link or edge e is associated to a wavefunction $\psi : SU(2) \to \mathbb{C}$, such that the wavefunction ψ belongs the two-dimensional square-integrable functions on $SU(2)$ which is itself a Hilbert space per link h_e

$$\psi \in L^2(SU(2)) \simeq h_e. \tag{354}$$

Recall that $SU(2)$ is diffeomorphic to S^3, which conjures memories and solutions of wavefunctions on a sphere: spherical harmonics.

The total Hilbert space of this quantum theory is a tensor product over all the edges e in the lattice E of individual Hilbert spaces h_e

$$\mathcal{H} = \otimes_{e \in E} = \otimes_{e \in E} L^2(SU(2)). \tag{355}$$

To build the Hamiltonian, introduce some oeprations on the Hilbert space. First, write the states as kets in the position basis

$$|\psi\rangle = \int_{SU(2)} dU \, \psi(U) \, |U\rangle \tag{356}$$

Where $|U\rangle$ is the position eigenvector defined by the three spatial coordinates. These states obey the inner product $\langle U|V\rangle = \delta(U - V)$.

Introduce the operators

$$L_U \, |\psi\rangle \equiv \int dV \, \psi(V) \, |UV\rangle \text{ and } R_U \, |\psi\rangle \equiv \int dV \, \psi(V) \, |VU^\dagger\rangle \, . \tag{357}$$

These operators are an analog of the shift operator $e^{ix\hat{p}}$. Note that $|UV\rangle$ is still a member of $SU(2)$. Differentiating these operators with respect to U will yield momentum operators: dynamics.

Lecture 15: Hamiltonian Lattice Gauge Theory

Continuing on the introduction to Hamiltonian lattice gauge theory as a means of quantization of gauge fields, we will build a microscopic formulation of gauge theory based on the real-space lattice. In contrast to the usual way of working on the Euclidean, Wick-rotated lattices, we will begin our theory with a Hamiltonian of classical degrees of freedom: namely, the parallel transporter U, a 2×2 matrix with determinant one, such that $U \in SU(2)$. Since we are working in 4D spacetime, we will have a 4D discretized lattice with lattice spacing $a \propto \frac{1}{K_c}$. In other words, 4D spacetime is discretized up to the cutoff K_c.

Like introduced before, each link, or edge, of the lattice $e \in E$, where E is the set of all links of the lattice, has an associated parallel transporter, corresponding to the shortest, rectilinear path in between each vertex, $U_e \in SU(2) \simeq S^3$. Note that the parallel transporter is not a local object, as it is path dependent, implicitly depending on more than one coordinate.

The classical configuration space for this lattice gauge theory is a Cartesian product of $SU(2)$ per link in the lattice

$$\mathcal{C} = SU(2) \times SU(2) \times \cdots \times SU(2). \tag{358}$$

To quantize this classical lattice gauge theory, we'll begin by taking the simplest guess possible and introduce a wavefunction $\psi : \mathcal{C} \to \mathcal{C}$, where the Hilbert space has a ket basis $|\psi\rangle \in \mathcal{H} \simeq \otimes_{e \in E} h_e$, where h_e is the space over each link.

Since $SU(2)$ is diffeomorphic to S^3, we are hinted towards defining wavefunctions as square-integrable functions on a sphere, and each point on the sphere will be associated to a complex number. Therefore, we define $h_e \equiv L^2(SU(2))$, where $L^2(SU(2))$ is an infinite-dimensional separable Hilbert space, since there are arbitrarily many orthogonal wavefunctions defined on the sphere.

Recall the two operators defined in $SU(2)$, for unitary matrix with unit determinant $U \in SU(2)$, that defined a right- and left-acting transformation

$$L_U : \ L^2(SU(2)) \to L^2(SU(2)) \tag{359}$$
$$R_U : \ L^2(SU(2)) \to L^2(SU(2)) \tag{360}$$

Where L_U and R_U commute, such that $[L_U, R_U] = 0$, and form the representation defined by the relations

$$L_U^\dagger L_U = R_U^\dagger R_U = \mathbb{I} \text{ and } L_{UV} = L_U L_V. \tag{361}$$

To understand how the infinite-dimensional Hilbert space $L^2(SU(2))$ breaks up into a direct sum of irreducible representations of $SU(2)$, we invoke the third part of the *Peter-Weyl therorem*, which states that the Hilbert space

over $SU(2)$ consisting of square-integrable functions may be regarded as a representation of a direct product of left- and right-acting operators, and the Hilbert space decomposes into an orthogonal direct sum of all the irreducible unitary representations, with multiplicity of each irreducible representation equal to its degree, the dimension of the underlying space of that representation (See Wikipedia page on Peter-Weyl theorem for overview). We write this all as

$$h_e \equiv L^2(SU(2)) \simeq \bigoplus_{l \in \frac{1}{2}\mathbb{Z}^+} V_l \otimes V_l^* \simeq \bigoplus_{l \in \frac{1}{2}\mathbb{Z}^+} \mathbb{C}^{2l+1} \otimes \mathbb{C}^{2l+1} \tag{362}$$

Where $V_l \simeq \mathbb{C}^{2l+1}$ is the $(2l+1)$-dimensional vector space furnishing the irreducible representation of $SU(2)$ of spin, or angular momentum l and $\frac{1}{2}\mathbb{Z}^+ = \{0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots\}$.

Now to find a representation of $SU(2)$ on this vector space V_l, we'll use a piece of Lie group representation theory not found in any textbook. Note that the action of $SU(2)$ generates a representation

$$\Pi_l(SU(2)) : \ V_l \to V_l. \tag{363}$$

In the procedure to calculate the matrix $\Pi_l(SU(2))$, we will use two pieces of machinery

- The spin-$\frac{1}{2}$ fundamental representation of $SU(2)$

  $$\Pi_{\frac{1}{2}}(U) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

- Tensor products of the spin-$\frac{1}{2}$ 2D vector spaces furnishing the fundamental representation of $SU(2)$, $V_{\frac{1}{2}} \simeq \mathbb{C}^2 \equiv \{|0\rangle, |1\rangle\}$.

Begin the procedure to get the matrix representation for any spin-l, take the tensor product of n of the spin-$\frac{1}{2}$ fundamental vector spaces

$$V_{\frac{1}{2}} \otimes V_{\frac{1}{2}} \otimes \cdots \otimes V_{\frac{1}{2}}. \tag{364}$$

Note that for quantum computer fans, this is the vector space of n qubits

$$V_{\frac{1}{2}} \otimes V_{\frac{1}{2}} \otimes \cdots \otimes V_{\frac{1}{2}} \simeq \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2 = \mathcal{C}^{2^n}. \tag{365}$$

The spin-l representation $\Pi_l(U)$, generated from $SU(2)$ action on V_l, lives in thsi tensor product space of n copies of $V_{\frac{1}{2}}$, as long as $n = 2l$ or $l = \frac{n}{2}$. Therefore, we will build V_l as a subspace of this tensor product space, most fo which will be thrown away once we find our subspace of interest, by building a set a $n+1$ orthonormal vectors

$$|w_{\frac{n}{2}}\rangle = |11\ldots1\rangle \tag{366}$$

$$|w_{\frac{n}{2}-1}\rangle = \frac{1}{\sqrt{n}}\left(|11\ldots10\rangle + |11\ldots101\rangle + \cdots + |01\ldots1\rangle\right) \tag{367}$$

$$\ldots \tag{368}$$

$$|w_{\frac{n}{2}-k}\rangle = \frac{1}{\sqrt{\binom{n}{k}}}\left(|1\ldots10\ldots0\rangle + (\text{all permutations of } k \text{ zeros and } n-k \text{ ones})\right) \tag{369}$$

$$\ldots \tag{370}$$

$$|w_{-\frac{n}{2}}\rangle = |00\ldots0\rangle\,. \tag{371}$$

Then the matrix elements of the representation $\Pi_l(U)$ are simply given by the expectation value on n copies of U in this orthonormal basis

$$[\Pi_l(U)]_{jk} = \langle w_j|\, U\otimes\cdots\otimes U\,|w_k\rangle \tag{372}$$

Where $j,\, k \in \{-\frac{n}{2},\ldots,\frac{n}{2}\}$.

Note that this method is good and fast for low spin representations, but clearly gets unwieldy for working with the Hilbert space of $n = 1000$ qubits, and the efficiency and value of the methods of addition of angular momentum and Clebsch-Gordan coefficients, with the raising and lower operators, the highest-weight vectors, etc.

Sticking to low spin representations for our purposes, we can start to extract matrix representations for $l = 0$, $l = \frac{1}{2}$, and $l = 1$.

For $l = 0$, the matrix representation is just the identity

$$\Pi_0(U) = \mathbb{I}. \tag{373}$$

For $l = \frac{1}{2}$, we have the fundamental matrix representation

$$[\Pi_{\frac{1}{2}}(U)]_{jk} = [U]_{jk}. \tag{374}$$

In the $l = 1$ subspace, there are three orthonormal basis vectors

$$\{|w_1\rangle = |11\rangle\,,\ |w_0\rangle = \frac{1}{\sqrt{2}}(|10\rangle + |01\rangle),\ |w_{-1}\rangle = |00\rangle\}. \tag{375}$$

The matrix elements are then gotten by the expectation value above, pulling values from the fundamental representation

$$[\Pi_1(U)]_{11} = \langle 11| U \otimes U |11\rangle = \langle 1| U |1\rangle \langle 1| U |1\rangle = a^2 \tag{376}$$

$$[\Pi_1(U)]_{00} = \frac{1}{2}\left(\langle 10| U \otimes U |10\rangle + \langle 10| U \otimes U |01\rangle + \langle 01| U \otimes U |10\rangle + \langle 01| U \otimes U |01\rangle\right) \tag{377}$$

$$= ad + bc. \tag{378}$$

Recall that the Peter-Weyl theorem tells us that the Hilbert space of square-integrable functions on $SU(2)$ is isomorphic to the infinite-dimensional, until we truncate, direct sum space of $(2l+1)$-dimensional vector spaces furnishing the representation of $SU(2)$

$$L^2(SU(2)) \simeq \bigoplus_{l \in \frac{1}{2}\mathbb{Z}^+} \mathbb{C}^{2l+1} \otimes \mathbb{C}^{2l+1}. \tag{379}$$

And $SU(2)$ acts on $L^2(SU(2))$ via the operator $L_U : L^2(SU(2)) \to L^2(SU(2))$ with action

$$L_U \simeq \bigoplus_{l \in \frac{1}{2}\mathbb{Z}^+} \Pi_l(U) \otimes \mathbb{I}. \tag{380}$$

The elements of the matrix representation $[\Pi_l(U)]_{jk} \equiv t^l_{jk}(U)$ are square-integrable functions from $SU(2)$ to the complex numbers $\mathbb{C}$, since $SU(2)$ is compact, and form an orthogonal (not orthonormal) basis for $L^2(SU(2))$, where $-l \leq j, k \leq l$. So, we can expand the wavefunction ket in the orthogonal basis of the matrix elements

$$|\psi\rangle = \sum_l \sum_{j,k} \psi^l_{jk} |j\rangle_l |k\rangle_l = \sum_l \sum_{j,k} \psi^l_{jk} \sqrt{2l+1}\, |t^l_{jk}\rangle\,. \tag{381}$$

The inner product of the Hilbert space $L^2(SU(2))$ is defined with the Haar measure $\int dU$ by

$$(\psi, \phi) \equiv \int dU\, \psi^*(U)\phi(U) \tag{382}$$

Where we use the inner product of the orthogonal basis vectors, where integrals over $SU(2)$ with the Haar measure behave like $\int dU\, U \otimes U^\dagger \propto \mathbb{I}\otimes\mathbb{I}$+(swap operations), and (**Exercise**)

$$(t^l_{jk}, t^{l'}_{j'k'}) = \delta^{ll'}\delta_{jj'}\delta_{kk'}\frac{1}{2l+1}. \tag{383}$$

Therefore, the basis for the total Hilbert space consists of wavefunctions $|\Psi\rangle \in \mathcal{H}$ and

$$|\Psi\rangle = \sum_{l_1 l_2 \ldots} \sum_{j_1 j_2 \ldots} \sum_{k_1 k_2 \ldots} \Psi^{l_1 l_2 \ldots}_{j_1 j_2 \ldots k_1 k_2 \ldots} |j_1\rangle_{l_1} |k_1\rangle_{l_1} |j_2\rangle_{l_2} |k_2\rangle_{l_2} \ldots. \tag{384}$$

To give dynamics to the Hilbert space, we define some observables. Consider an element of $SU(2)$, $U = e^{c_j \tau^j}$, where $c_j \tau^j$ are elements of the Lie algebra of $SU(2)$, $c_j \in \mathbb{R}^3$, and τ^j are the Pauli spin matrices multiplied by i and divided by 2 for normalization conditions. The spin matrices obey the commuation relations $[\tau^j, \tau^k] = -2\epsilon^{jk}{}_l \tau^l$. We can recover the spin matrix from the group element via differentation

$$\frac{dU}{dc_j}\Big|_{c_j=0} = \tau^j. \tag{385}$$

The first observable we define is the anti-Hermitian *left angular momentum operator*

$$\hat{l}_L^j \equiv \frac{d}{ds} L_{U=e^{s\tau^j}}\Big|_{s=0}. \tag{386}$$

Note that this is a factor of i away from being Hermitian, and is analogous to the Hermitian linear momentum operator $i\hat{p}$. Also, notice that we begin with a unitary operator L_U, apply an anti-Hermitian operator $\frac{d}{ds}$, and kill off the unitarity by setting $s = 0$ after the derivative.

The second observable we define, similar to the first, is the *right angular momentum operator*

$$\hat{l}_R^j \equiv \frac{d}{ds} R_{U=e^{s\tau^j}}\Big|_{s=0}. \tag{387}$$

Next, define the position observable $\hat{U}_{jk}$, which is also a map $L^2(SU(2))rightarrowL^2(SU(2))$ and yields the fundamental representation matrix elements when it acts on the position eigenkets $|U\rangle$ of $SU(2)$.

$$\hat{U}_{jk} |U\rangle = [\Pi_{\frac{1}{2}}(U)]_{jk} |U\rangle . \tag{388}$$

Acting the position operator on the Hilbert space wavefunctions $|\psi\rangle = \int dU\, \psi(U) |U\rangle$, we get the Haar integral over position eigenkets

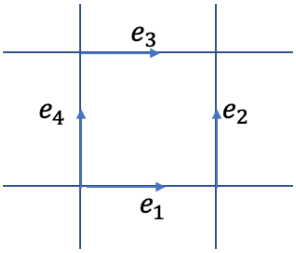
$$\hat{U}_{jk} |\psi\rangle = \int dU\, \psi(U) [\Pi_{\frac{1}{2}}(U)]_{jk} |U\rangle = \int dU\, \psi(U) t_{jk}^{l=\frac{1}{2}}(U) |U\rangle . \tag{389}$$

To build the Hamiltonian, kinetic energy plus potential energy, we use the plaquette operator to define parallel transport on the 4D lattice

$$\begin{aligned} \hat{U}_\square : & L^2(SU(2)) \otimes L^2(SU(2)) \otimes L^2(SU(2)) \otimes L^2(SU(2)) \\ & \to L^2(SU(2)) \otimes L^2(SU(2)) \otimes L^2(SU(2)) \otimes L^2(SU(2)) \end{aligned}$$

We choose the convention to set arrows on each link running "left-to-right" and "down-to-up" in the plane of the "paper", and then walk around the plaquette counterclockwise (CCW), taking the Hermitian conjugate of the parallel transporter

if we are traveling against the arrow. The parallel transporter for each link is considered the observable for that link when traveling around the plaquette. Each link are labeled by e_i, $i = 1, 2, 3, 4$.



Build an operator that acts on the 4D space of links, as a sum over the links of the direct products of parallel transporters

$$\hat{M}_{\square;j_1,k_4} = \sum_{k_1,k_2,k_3} \hat{U}_{j_1k_1} \otimes \hat{U}_{k_1k_2} \otimes \hat{U}^{\dagger}_{k_2k_3} \otimes \hat{U}^{\dagger}_{k_3k_4}. \tag{390}$$

This is then turned into an observable by summing over the last, initially-fixed index k_4

$$tr(\hat{U}_{\square}) \equiv \sum_{k_4} \hat{M}_{\square;k_4,k_4}. \tag{391}$$

Note that $tr(\hat{U}_{\square})$ is a trace operator that acts on $L^2(SU(2)) \otimes L^2(SU(2)) \otimes L^2(SU(2)) \otimes L^2(SU(2))$, and **not** a number!

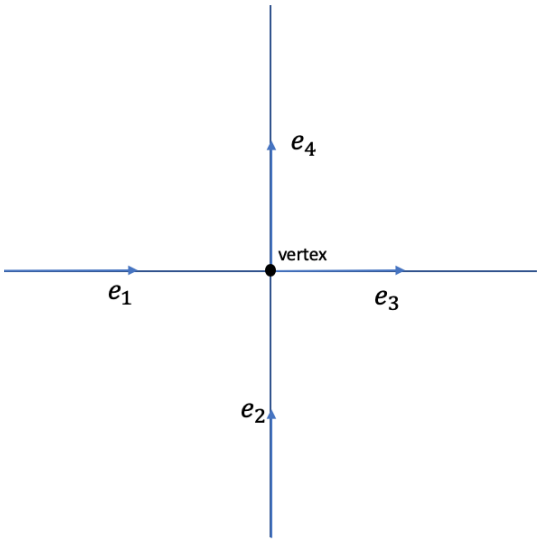
Put the defined observables together into the *Kogut-Susskind Hamiltonian*

$$\hat{H}_{KS} = -\frac{g_H^2}{2a} \sum_{e \in E} \hat{l}^j_L(e)\hat{l}^j_L(e) + \frac{1}{2g_H^2 a} \sum_{\text{plaquettes } \square} \Big(tr(\hat{U}_{\square}) + \text{Hermitian conjugate} \Big) \tag{392}$$

Where g_H is the coupling constant.

This model has a huge group of gauge symmetries.

Consider a vertex in the lattice and build the operator



$$M_x \equiv R_x(e_1) \otimes R_x(e_2) \otimes L_x(e_3) \otimes L_x(e_4) \tag{393}$$

Where $x \in SU(2)$. These operators obey the following commutation relations

$$[M_x(v), M_y(w)] = 0 \text{ and } [M_x(v), \hat{H}_K S] = 0, \text{ for all } x, y, v, w. \tag{394}$$

In summary, Wilson's formulation received much more attention at the time for its ease of discretizstion, translation to computer programs, and use of Monte Carlo sampling, which classical computers are good at. Kogut and Susskind argued that the plaquette operator becomes the curvature term, as expected, in the small lattice spacing a limit in their theory, as well as that hte kinetic energy term becomes the kinetic energy in the timelike direction of the curvature term. The Kogut-Susskind formulation may be very promising for quantum simulations done by quantum computers, which are not very good at sampling techniques, but excel in simulating the dynamics of local lattice models.

Lecture 16: Spontaneous Symmetry Breaking

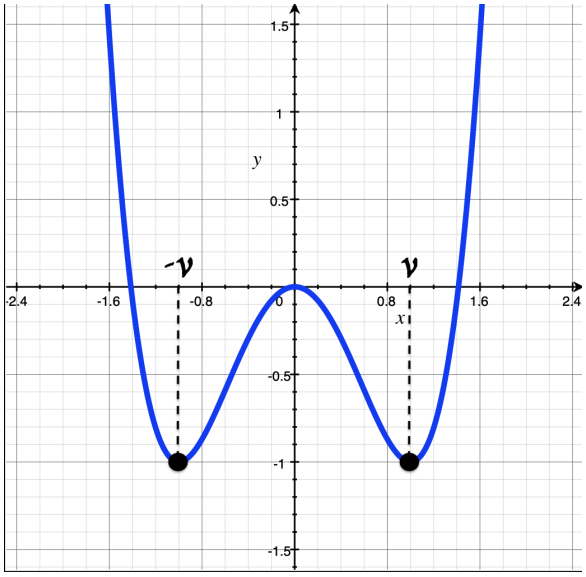
This is the last topic that we need to complete our description of the standard model of elementary particle interactions. *Spontaneous symmetry breaking* (SSB) is an observed behavior within field theories, and we will begin with some examples of classical SSB.

Particle in a double well example

Consider a classical particle confined to a double-well potential with Hamiltonian

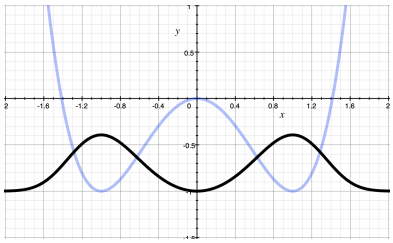
$$H = \frac{p^2}{2m} + V(x). \tag{395}$$

This system posses $\mathbb{Z}_2$ symmetry in its solutions, since $x \to -x$ is a symmetry operation. Note that the particle in this potential at $x = 0$ must choose a positive or negative local minimum, meaning that the ground state is degenerate and breaks the $\mathbb{Z}_2$ symmetry.

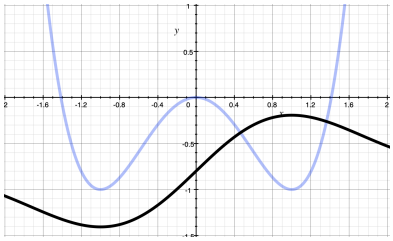


Potential exhibiting $\mathbb{Z}_2$ symmetry.

The quantum analog of this classical theory does not exhibit SSB, since the ground state wavefunction is symmetric, but the first excited state is antisymmetric.



Symmetric ground state wavefunction.



Antisymmetric first excited state wavefunction.

Ising model (statistical physics) example

The Ising model is a model of ferromagnetic materials, where spins can point up or down, corresponding to a spin value $s_j = \pm 1$, with a Hamiltonian containing the pairwise summation

$$H = -\sum_{\langle jk \rangle} s_j s_k. \tag{396}$$

This theory contains two ground states: all spins pointing up or all pointing down, and possesses $\mathbb{Z}_2$ symmetry by the operation $s_j \to -s_j$.

Depending on the temperature, the thermal state of this system can be in one of two regimes: critical or non-critical. Consider the Gibbs, or mixed, state density operator describing the system at any temperature

$$\rho = \frac{e^{-\beta H}}{Z} \tag{397}$$

Where Z is the partition function and $\beta = \frac{1}{k_B T}$ is the temperature factor. Note that the Gibbs state is all $\mathbb{Z}_2$ symmetric, meaning that it does break any symmetries. Now there exists a critical temperature β_c, below which the system becomes ordered due to small external magnetic fluctuations and symmetry is broken. Above the critical temperature, the system is disordered with random thermal fluctuations. As $\beta \to \infty$,

$$\rho = \left(\frac{1}{2} - \epsilon\right) (\text{all up states}) + \left(\frac{1}{2} - \epsilon\right) (\text{all down states}). \tag{398}$$

Classical field theory example

Consider the Lagrangian density, much like the ϕ^4 interacting theory we've encountered, but the "mass" term is made negative and $m \to \mu$

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) + \frac{1}{2}\mu^2\phi^2 - \frac{\lambda}{4!}\phi^4. \tag{399}$$

Note that measurable quantities are usually functions of the parameters, making this a perfectly fine theory. We then have the Hamiltonian

$$H = \int d^3x \left(\frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 - \frac{1}{2}\mu^2\phi^2 + \frac{\lambda}{4!}\phi^4\right). \tag{400}$$

To uncover the $\mathbb{Z}_2$ symmetry of this theory, minimize the Hamiltonian with respect to the field ϕ, making all derivatives of ϕ equal to zero, and calculate the configuration with the smallest energy. The extremum condition for the remaining potential energy terms

$$\frac{\partial V(\phi)}{\partial \phi} = -\mu^2\phi + \frac{\lambda}{6}\phi^3 = 0 \tag{401}$$

Yields three configurations that extremize the energy of the system. Namely, for $\phi = 0$ and $\phi = \pm\sqrt{\frac{6\mu^2}{\lambda}}$. The quantum analog of this theory also exhibits SSB.

Quantum field theory example: transverse Ising model

Consider the quantum theory of a 1D lattice of spins, which also exhibits $\mathbb{Z}_2$ symmetry, with the Hamiltonian

$$\hat{H} = -\sum_j \sigma_j^x \sigma_{j+1}^x + h\sum_j \sigma_j^z \tag{402}$$

Where the first term is the neighboring interaction and the second term is the magnetic interaction with h as the magnetic field strength, and the Pauli spin matrices

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ and } \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{403}$$

The Hamiltonian exhibits $\mathbb{Z}_2$ symmetry, since

$$[\Phi, \hat{H}] = 0 \text{ for } \Phi = \ldots \sigma_j^z \sigma_{j+1}^z \ldots. \tag{404}$$

The basis for the $h = 0$ ground state can be written as

$$|\Omega_+\rangle \equiv |+\rangle \otimes |+\rangle \otimes \cdots \otimes |+\rangle \tag{405}$$
$$|\Omega_-\rangle \equiv |-\rangle \otimes |-\rangle \otimes \ldots |-\rangle \tag{406}$$

Where the individual spin states are

$$|+\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) \text{ and } |-\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right). \tag{407}$$

The $h = 0$ ground state eigenspace is twofold degenerate with the above basis. A good, non-degenerate ground state, for example, could be $\frac{1}{\sqrt{2}}(|\Omega_+\rangle \pm |\Omega_-\rangle)$, but these states are never seen in experiment, since decoherence destroys any superpositions of states. Neither state by itself exhibits the symmetry, but one state must be chosen by measurement. The only information we have about the two states' relationship is $\Phi\,|\Omega_+\rangle = |\Omega_-\rangle$.

Continuous SSB Example: Linear σ-model

Consider the Lagrangian density for an effective model of pions that exhibits SSB of a continuous symmetry

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi^j)^2 + \frac{1}{2}\mu^2(\phi^j)^2 - \frac{\lambda}{4!}(\phi^j)^4 \tag{408}$$

Where $(\phi^j)^2 = \sum_{j=1}^{N} \phi^j \phi^j$. The dynamics of the N independent (Klein-Gordon) scalar field are invariant under orthogonal rotations $O \in O(N)$, a continuous group of symmetries, such that

$$O: \ \phi^j \to [O]^j_k \phi^k. \tag{409}$$

What is the lowest energy configuration? Minimize the potential energy $V(\phi^j) = -\frac{1}{2}\mu^2(\phi^j)^2 + \frac{\lambda}{4!}(\phi^j)^4$ with respect to ϕ^j, showing that the minimum occurs for any constant configuration of the fields that satisfies the equation (**Exercise**)

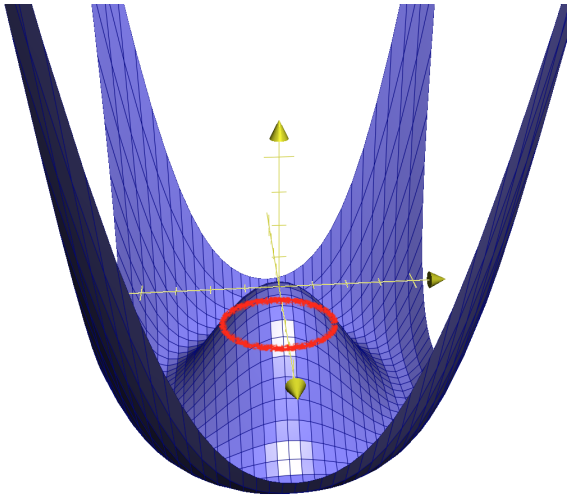
$$(\phi_0^j)^2 = \frac{\mu^2}{\lambda}. \tag{410}$$

There is more than one energy configuration that leads to this solution, including superpositions of the following vectors

$$\begin{pmatrix} \frac{\mu}{\sqrt{\lambda}} \\ 0 \\ 0 \\ \cdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \frac{\mu}{\sqrt{\lambda}} \\ 0 \\ \cdots \\ 0 \end{pmatrix}, \ldots . \tag{411}$$

In the case of $N = 2$, the potential energy minima form the "wine bottle" or "mexican hat" potential. The minimum energy configurations are points on this potential and form circles around the indent of the "bottom of the bottle". In

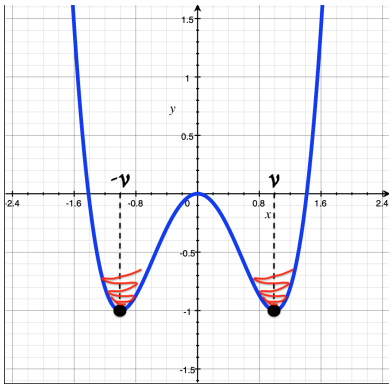
other words, any configuration that lands on the circle is a minimum energy configuration.



Mexican hat potential of the linear sigma model.

Low Energy Dynamics for SSB

To study the dynamics as we start to depart from the minimal energy configurations, the ground state, consider the $\mathbb{Z}_2$-symmetry case. Effectively, the low energy dynamics for a classical $\mathbb{Z}_2$-symmetric system are that of a harmonic oscillator with an effective mass, a restoring force.



Small fluctuations in the energy behave much like the harmonic oscillator.

In the continuous case, we choose coordinates by rotating such that

$$\phi_0^j = \begin{pmatrix} 0 \\ 0 \\ \dots \\ \nu \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \dots \\ \frac{\mu}{\sqrt{\lambda}} \end{pmatrix} \tag{412}$$

And see how this behaves with small energy fluctuations. Define *shifted fields* in terms of some new coordinates where the vector of fields $\underline{\phi}$ is now defined as

$$\underline{\phi}(x) \equiv (\pi^k(x), \nu + \sigma(x)), \text{ where } k = 1, \dots, N-1. \tag{413}$$

Note that π is not the conjugate momentum density, but a new classical field, which will be, hence the notation, the pion, and k denotes the vector index. Rewrite $\mathcal{L}$ in terms of the shifted fields (**Exercise**)

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \pi^k)^2 + \frac{1}{2}(\partial_\mu \sigma)^2 \tag{414}$$

$$- \frac{1}{2}\left(2\mu^2\sigma^2 - \sqrt{\lambda}\mu\sigma^3 - \sqrt{\lambda}\mu(\pi^k)^2\sigma - \frac{\lambda}{4}\sigma^4 - \frac{\lambda}{2}(\pi^k)^2\sigma^2 - \frac{\lambda}{4}(\pi^k)^4\right). \tag{415}$$

There are $N-1$ massless π^k fields and one massive σ field. The second and third terms in $\mathcal{L}$ correspond to an effective massive Klein-Gordon scalar field. The $N-1$ π^k fields are effectively massless, as all of the other terms above contain λ and are interaction terms.

It costs energy to move transverse in the potential, perpendicular to the circle of minima, corresponding to the effective mass of the σ field. To move tangentially to the manifold of minima, the circle, it costs no energy, corresponding to the massless π fields.

Goldstone's Theorem

In the $O(N)$ linear σ-model there are $\binom{N}{2}$ independent continuous symmetries, the dimension of the rotation group $O(N)$. After SSB, there are $\binom{N-1}{2}$ remaining symmetries, the dimension of $O(N-1)$, corresponding to rotations of the π^k fields. The number of broken symmetries is equal to $\binom{N}{2} - \binom{N-1}{2} = (N-1)$, which is also the number of massless fields. In other words, each broken symmetry causes a massless excitation: the *Goldstone modes* or *Goldstone bosons*.

Therorem: For every broken symmetry, there is a corresponding massless bosonic particle.

Proof:

Consider a classical field theory with fields $\phi^a(x)$, $a = 1, 2, \ldots$, and the general Lagrangian density $\mathcal{L} = (\text{derivatives}) - V(\phi^a)$.

Let ϕ_0^a be a constant (in an extrema) field that minimizes the potential such that

$$\frac{\partial V(\phi^a)}{\partial \phi^a}\Big|_{\phi^a = \phi_0^a} = 0. \tag{416}$$

Then expand the potential, a function of the vector of fields $\phi_0^a \equiv \underline{\phi}_0$, near the minima

$$V(\phi^a) = V(\phi_0^a) + \frac{1}{2}(\phi^a - \phi_0^a)(\phi^b - \phi_0^b)\left(\frac{\partial^2 V}{\partial \phi^a \partial \phi^b}\right)\Big|_{\phi = \phi_0}. \tag{417}$$

Call the Hessian matrix $[m^2]_{ab}$, which is symmetric and real and the eigenvalues give the masses of the effective fields.

With an orthogonal rotation, we can diagonalize a symmetric, real matrix $m^2 \to O^T D O$. Redefine the fields $\pi^a = [O]_b^a \phi^b$, and rewrite the Lagrangian density

$$\mathcal{L}(\pi^a) = (\text{derivatives}) - \sum_a D_a^2 (\pi^a)^2 \tag{418}$$

Where the eigenvalues correspond to the masses of the π particles, and we must now show that there are eigenvalues equal to zero. In other words, every continuous symmetry leads to an eigenvalue equal to zero.

A general, global, continuous symmetry of the fields has the form

$$\phi^a \to \phi^a + \alpha \Delta^a(\phi) \tag{419}$$

Where α is infinitesimal and Δ^a is a shift function of the fields. This is a symmetry of the potential, since it causes the derivatives to vanish such that

$$V(\phi^a) = V(\phi^a + \alpha \Delta^a(\phi)). \tag{420}$$

This implies, by expanding and equating first order terms, the directional derivative of the potential is zero

$$\Delta^a(\phi) \frac{\partial}{\partial \phi^a} V(\phi) = 0. \tag{421}$$

Differentiate this with respect to ϕ^b to get

$$\frac{\partial \Delta^a(\phi)}{\partial \phi^b} \frac{\partial V(\phi)}{\partial \phi^a} + \Delta^a(\phi) \frac{\partial^2 V}{\partial \phi^a \partial \phi^b} = 0. \tag{422}$$

Evaluate at $\phi^a = \phi_0^a$ to get

$$\sum_a \Delta^a(\phi_0^a)[m^2]_{ab} = 0 \tag{423}$$

Where $\sum_a \Delta^a(\phi_0^a) = \Delta^T$ is the zero eigenvector, where the $\Delta^a(\phi_0^a)$ are linearly independent for each continuous symmetry, which follows by definition of the general, global, continuous symmetry imposed above.

Lecture 17: