QUANTUM FIELD THEORY

Mentee: Mahadevan Subramanian **Mentor: Som Phene**

Contents

ABSTRACT

This is a report on my explorations in learning Quantum Field Theory in SoS 2021.

1 Some Preliminaries

1.1 A brief idea for closeness without distance

Definition 1.1 (Convergence)**.** For a given metric space $\mathcal{M} = (X, d)$ and a sequence $\{x_n\}_{n\in\mathbb{N}}$ we say x_n converges to x iff $\forall\epsilon > 0, \exists N_0$ such that $\forall n > N_0, d(x_n, x) < \epsilon$.

Here we must use a metric space with a distance metric d to quantify closeness. However, it is not necessary that we may be able to construct such a metric for any space and so we need to generalize a little more.
Let us define a set $\mathcal{T}$ where each element is a set of points from a space X and it satisfies the following properties.

- $X, \emptyset \in \mathcal{T}$
- $O_1, \dots O_n \in \mathcal{T}$ implies $O_1 \cap \cdots \cap O_n \in \mathcal{T}$ (finite intersections only)
- $O_i, i \in I \in \mathcal{T}$ implies $\bigcup_i O_i \in \mathcal{T}$ (any unions)

This set $\mathcal{T}$ will contain all the open sets of X. Now while we have an ϵ, δ definition of continuity we can obtain a nicer definition of continuity from the following theorem.

Theorem 1.1 (Open-mapping theorem)**.** For a continuous function $f : X \to Y$, the pre image of any open set in Y would be an open set in X.

This directly gives us a very nice definition for continuity which does not invoke the requirement for a metric

1.2 Special Theory of Relativity revision

Transformation laws

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu} \tag{1}$$
$$g_{\mu\nu} = \Lambda^{\rho}{}_{\mu} g_{\rho\lambda} \Lambda^{\lambda}{}_{\nu} \tag{2}$$
$$g_{\mu\nu} = \text{diag}(1, -1, -1, -1) \tag{3}$$
$$\tag{4}$$

Maxwell's laws

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{5}$$
$$\partial_\mu F^{\mu\nu} = 4\pi j^\mu \tag{6}$$
$$\partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} + \partial_\rho F_{\mu\nu} = 0 \tag{7}$$

Lorentz force

$$\frac{\partial p^\mu}{\partial \tau} = q F^{\mu\nu} u_\nu \tag{8}$$

Energy momentum tensor

$$T^{\mu\nu} = \sum_n \int d\tau p_n^\mu p_n^\nu \delta^4(x^\rho(\tau) - x_n^\rho(\tau)) \tag{9}$$

$$T^{\mu\nu} = -pg^{\mu\nu} + (p+\rho)u^\mu u^\nu \tag{10}$$

Source flux theorem

$$\partial_\mu T^{\mu\nu} = 0 \tag{11}$$

1.3 Quantum Mechanics revision

Here I will just list some very useful formulae of quantum mechanics that I am so far familiar.

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(x)\right]\psi(\vec{r},t) = \iota\hbar\frac{\partial\psi(\vec{r},t)}{\partial t} \tag{12}$$

$$\lim_{\epsilon\to 0}\left(-\frac{\hbar^2}{2m}\left[\frac{d\psi(x+\epsilon)}{dx} - \frac{d\psi(x-\epsilon)}{dx}\right] + \int_{x-\epsilon}^{x+\epsilon} V\psi(x)dx\right) = 0 \tag{13}$$

$$[\hat{x}_i, \hat{p}_j] = \iota\hbar\delta_{ij} \tag{14}$$

$$\frac{d}{dt}\langle\hat{A}\rangle = \frac{1}{\iota\hbar}\langle[\hat{A},\hat{H}]\rangle + \left\langle\frac{d\hat{A}}{dt}\right\rangle \tag{15}$$

Dual of vector $\hat{O}\left|\xi\right\rangle$ is $\left\langle\xi\right|\hat{O}^\dagger$ and given an orthogonal basis $\{\left|\phi_i\right\rangle\}$ we have

$$\sum_i \left|\phi_i\right\rangle\left\langle\phi_i\right| = \mathbb{I} \tag{16}$$

All observables must be hermitian, for example radial momentum is given by

$$\hat{p}_r = \frac{\hat{r}\cdot\vec{p} + \vec{p}\cdot\hat{r}}{2} \tag{17}$$

Spherical harmonics

$$Y_m^l(\theta,\phi) = \Theta(\theta)\Phi(\phi) \tag{18}$$

$$sin(\theta)\frac{d}{d\theta}\left(sin(\theta)\frac{d\Theta(\theta)}{d\theta}\right) + [l(l+1)sin^2(\theta) - m^2]\Theta(\theta) = 0 \tag{19}$$

$$\frac{d^2\Phi(\phi)}{d\phi^2} = -m^2\Phi(\phi) \tag{20}$$

Angular momentum algebra

$$L_i = \epsilon_{ijk}x_j p_k \tag{21}$$

$$[L_i, L_j] = \epsilon_{ijk}\iota\hbar L_k \tag{22}$$

Harmonic oscillator ladder operations

$$[\hat{a}, \hat{a}^\dagger] = \mathbb{I} \tag{23}$$

$$\hat{H} = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) \tag{24}$$

$$\hat{a}\left|n\right\rangle = \sqrt{n}\left|n-1\right\rangle \tag{25}$$

$$\hat{a}^\dagger\left|n\right\rangle = \sqrt{n+1}\left|n+1\right\rangle \tag{26}$$

$$\left|\alpha\right\rangle = e^{-|\alpha|^2/2}D(\alpha)\left|0\right\rangle,\ D(\alpha) = e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}} \tag{27}$$

Addition of angular momentum

$$J_i = L_i + S_i \tag{28}$$

$$[J_i, J_j] = \iota\hbar\epsilon_{ijk}J_k \tag{29}$$

$$J_\pm = J_1 \pm J_2,\ [J_+, J_-] = 2\hbar J_3 \tag{30}$$

$$J^2\left|j,m\right\rangle = j(j+1)\hbar^2\left|j,m\right\rangle,\ J_3\left|j,m\right\rangle = m\hbar\left|j,m\right\rangle \tag{31}$$

$$J_\pm\left|j,m\right\rangle = \hbar\sqrt{(k\mp q)(k\pm q+1)}\left|j,m\pm 1\right\rangle \tag{32}$$

Rotations in a space

$$U_R(\hat{n},\theta) = \exp\left(-\frac{\iota J\cdot\hat{n}}{\hbar}\theta\right) \tag{33}$$

Spherical tensors

$$T_q^{(k)} = Y_{l=k}^{m=q}(T) \tag{34}$$

$$[J_z, T_q^{(k)}] = \hbar q T_q^{(k)} \tag{35}$$

$$[J_\pm, T_q^{(k)}] = \hbar\sqrt{(k\mp q)(k\pm q+1)}T_{q\pm1}^{(k)} \tag{36}$$

Wigner-Eckhart Theorem

$$\langle j', m'|\, T_q^{(k)}\, |j,m\rangle = \langle j'||T^{(k)}||j\rangle\langle j,k;m,q|j'm'\rangle \tag{37}$$

Here $\langle j,k;m,q|j'm'\rangle$ is the CG coefficient for the angular momenta addition of $|j,m\rangle$ and $|k,q\rangle$ to get $|j',m'\rangle$

2 Calculus of Variations

The main reference for this section is [14].

2.1 Introduction

The calculus of variations concerns itself with the extremals of functions whose domain is the space of curves (hence being infinite dimensional). Such functions are referred to as functionals. An example of one such functional is $\Phi(\gamma) = \int_{t_0}^{t_1}\sqrt{1+\dot{x}^2}dt$ where $\gamma = \{(t,x) : x(t) = x, t_0 \le t \le t_1\}$. An approximation γ' to γ is simply $\gamma' = \{(t,x) : x = x(t) + h(t), t_0 \le t \le t_1\}$

Definition 2.1 (Differentiable functional)**.** Φ is a differentiable functional if $\Phi(\gamma+h) - \Phi(\gamma) = F + R$ where F is linear in h and R is $\mathcal{O}(h^2)$. The linear part $F(h)$ is referred to as the differential.

The differential of a function is also called its variation and if it exists, it is unique for the given functional. For example the functional $\Phi(\gamma) = \int_{t_0}^{t_1} L(x,\dot{x},t)$ the differential can be reached by this definition to be equal to

$$F(h) = \int_{t_0}^{t_1}\left[\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}}\right]hdt + \left(\frac{\partial L}{\partial \dot{x}}h\right)\Bigg|_{t_0}^{t_1} \tag{38}$$

The proof of this follows from expressing $L(x+h,\dot{x}+\dot{h},t)$ in it's multi-variable Taylor's expansion from $L(x,\dot{x},t)$. It follows that seeing first order only we get $\Phi(\gamma+h) - \Phi(\gamma) = \int\left[\frac{\partial L}{\partial x}h + \frac{\partial L}{\partial \dot{x}}\dot{h}\right] + \mathcal{O}(h^2)$ which simplifies using integration by parts to the expression in equation 38.

Definition 2.2 (Extremal)**.** An extremal of a differentiable functional $\Phi(\gamma)$ is the curve γ where $F(h) = 0$ for all h.

We can see from equation 38 that under certain boundary conditions of taking $x(t_0) = x_0$ and $x(t_1) = x_1$, the extremal condition arises for the curve $\gamma : x = x(t)$ if and only if

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = 0 \tag{39}$$

This is exactly the Euler-Lagrange equation which shows that it follows from principle of least action. If we take x to be a n dimensional vector in $\mathbb{R}^n$, we get γ to be a $n+1$ dimensional curve in $\mathbb{R}\times\mathbb{R}^n$ and similarly L is a function in $2n+1$ variables. Hence the extremal now comes to be n second order equations with the $2n$ conditions from $x(t_0) = x_0$ and $x(t_1) = t_1$.

The extremal for the functional with $L(x,\dot{x},t) = \sqrt{1+\dot{x}^2}$ comes out to be a straight line connecting the two points in the (x,t) space. Similarly if one were to find the extremal for distance along a spherical surface it comes to be the great circle connecting the two points. This is generalized as geodesic equations and are used for the equations of motion in general relativity. Using the function $L(x,\dot{x},t) = \sqrt{g_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta}$ the equations of motion arise to be

$$\frac{d^2x^\alpha}{d\tau^2} + \Gamma^\alpha_{\beta\gamma}\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} = 0 \tag{40}$$

Here $\Gamma^\alpha_{\beta\gamma} = g^{\alpha\mu}(g_{\beta\mu,\gamma} + g_{\gamma\mu,\beta} - g_{\beta\gamma,\mu})$ which is the affine connection. This equation motion does infact arise from Hamilton's principle of least action.

Definition 2.3 (Lagrangian)**.** We define $L(q, \dot{q}, t) = T - U$ as the lagrangian where q_i are the generalized coordinates and $\dot{q}_i$ are the generalized velocities. We define $p_i = \partial L/\partial \dot{q}_i$ as the generalized momenta and $\partial L/\partial q_i$ as the generalized force. The extremal of $\int_{t_0}^{t_1} L dt$ which is the action is arrived at the solution of the Euler-Lagrange equations.

2.2 The Legendre Transformation

Let $f : \mathbb{I} \to \mathbb{R}$, $\mathbb{I} \subseteq \mathbb{R}$ be a convex function. The legendre transformation is a new function $g : \mathbb{I}^* \to \mathbb{R}$, $\mathbb{I}^* \subseteq \mathbb{R}$ defined as follows

$$g(p) = \sup_{x \in \mathbb{I}} (xp - f(x)) \tag{41}$$

The farthest vertical distance between the line $y = px$ and $y = f(x)$ comes out to be the expression of $g(p)$ which is exactly it's definition. More solidly, let $F(x, p) = px - f(x)$, and $F(x, p)$ has a maximum w.r.t x at some $x(p)$, then $g(p) = F(p, x(p))$. This extremal occurs at $f'(x) = p$ hence we get the expression in equation 41.
If the function f is convex over some domain we know that it's derivative will have an inverse over that same domain hence guaranteeing that $x(p)$ must be unique if it exists. Also if for a value of p the supremum is unbounded, that cannot be in $\mathbb{I}^*$.
An example for this is the function $f = cx$ which is convex but not strictly convex. If $p \neq c$ the supremum of $F(x, p)$ is non finite hence the domain is only $\{c\}$ and $g(c) = 0$. We can similarly find that for $f(x) = mx^2/2$, $g(p) = p^2/2m$. More generally for $f(x) = x^\alpha/\alpha$ we get $g(p) = p^\beta/\beta$ for $\alpha + \beta = \alpha\beta$ and both $\alpha, \beta > 1$.
Importantly the Legendre transform is involvutive hence it's square is the identity. This becomes readily apparent on writing the transform in terms of derivatives as shown below.

$$g(p) = p(f')^{-1}(p) - f((f')^{-1}(p)) \tag{42}$$

This form of writing the transform is not entirely correct since it brings up issues for non convex functions.

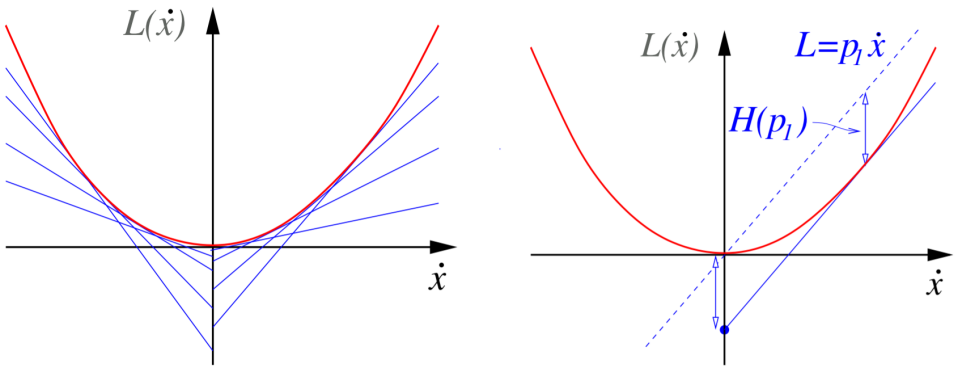

Figure 1: Left panel: envelope of the family of lines would give back the Legendre transform of the intercept value of a certain slope. Right panel: the supremum of the $p\dot{x} - L(\dot{x})$ can be geometrically represented as the y intercept of the tangent of same slope. Image credit: [3].

Corollary 2.0.1. Consider a given family of straight lines $y = px - g(p)$, the envelope of these lines has the equation $y = f(x)$ where $f(x)$ is the Legendre transformation of g

This above corollary follows from the fact that the enveloping function for a family of lines has these lines to be it's tangents. Hence this gives $p = \partial f/\partial x$ and at this value $g(p) = \frac{\partial f}{\partial x} x - f(x)$ which is the supremum of the expression $px - f(x)$ and hence g and p are Legendre transforms of each other.

Definition 2.4. Two functions, f and g, which are Legendre transforms of each other are called *dual in the sense of Young*.

Since $g(p)$ is defined as the supremum over x of $px - f(x)$, we get the following inequatlity referred to as the *Young's inequality*

$$px \leq f(x) + g(p) \tag{43}$$

Now let's suppose we have some convex function $f(\mathbf{x})$ where $\mathbf{x} = (x_1, \ldots, x_n)$ and we now define a $\mathbf{p} = (p_1, \ldots, p_n)$. We define the legendre transform as follows

$$g(\mathbf{p}) = \sup_{\mathbf{x}} \left(\langle \mathbf{p}, \mathbf{x} \rangle - f(\mathbf{x}) \right) \tag{44}$$

Since one would reach an maxima when the gradient of the expression is zero, this gives us the condition $\mathbf{p} = \partial f / \partial \mathbf{x}$. If we take $f : \mathbb{R}^n \to \mathbb{R}$ as a convex function and denote the dual space by $\mathbb{R}^{n*}$, we essentially have $f' : \mathbb{R}^n \to \mathbb{R}^{n*}$ where $f'(\mathbf{x}) = \partial f / \partial \mathbf{x}$. The $\mathbf{p}$ are in the dual vector space.

Definition 2.5 (Convex function)**.** A multi-variable function is convex over a convex set iff it's Hessian matrix of second derivatives is positive semidefinite over this convex set.

The hessian matrix is defined as $(\mathbf{H_f})_{i,j} = \dfrac{\partial^2 f}{\partial x_i \partial x_j}$. We must note that this Hessian matrix is the Jacobian for the function $\partial f / \partial \mathbf{x}$. Using the inverse function theorem, a positive semidefinite Hessian matrix implies that the derivative function has an invertible Jacobian and since this is over the whole domain of $\mathbb{R}^n$, the mapping of $g : \mathbb{R}^{n*} \to \mathbb{R}^n$ is well defined.
As an example we can pick the function $f(\mathbf{x}) = f_{ij} x_i x_j$ (repeated indices are summed over). The gradient of this function is $f'(\mathbf{x})_i = (f_{ki} + f_{ik}) x_k$ so putting in the expression we get $g(\mathbf{p}) = x_i (f_{ki} + f_{ik}) x_k - x_i f_{ik} x_k = f_{ij} x_j x_i$. Since $x_i x_j = x_j x_i$ this shows that $g(\mathbf{p}(\mathbf{x})) = f(\mathbf{x})$ and since these are Legendre transforms of each other, we also have $f(\mathbf{x}(\mathbf{p})) = g(\mathbf{p})$.

2.3 Hamilton's mechanics

Theorem 2.1. The system of Lagrange's equations are equivalent to the system of $2n$ first order equations which are referred to as the Hamilton's equations.

$$\dot{\mathbf{p}} = -\frac{\partial \mathcal{H}}{\partial \mathbf{q}} \tag{45}$$

$$\dot{\mathbf{q}} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}} \tag{46}$$

Here $\mathcal{H}(\mathbf{p}, \mathbf{q}, t) = \langle \mathbf{p}, \dot{\mathbf{q}} \rangle - L(\mathbf{q}, \dot{\mathbf{q}}, t)$ which is the Legendre transformation of the lagrangian with respect to $\dot{\mathbf{q}}$.

Proof. These equations arise from writing the differential of $\mathcal{H}$ which can expressed simply by chain rule and also in terms of the Legendre transform of L hence

$$d\mathcal{H} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}} \cdot d\mathbf{p} + \frac{\partial \mathcal{H}}{\partial \mathbf{q}} \cdot d\mathbf{q} + \frac{\partial \mathcal{H}}{\partial t} dt = \dot{\mathbf{q}} \cdot d\mathbf{p} - \frac{\partial L}{\partial \mathbf{q}} \cdot d\mathbf{q} - \frac{\partial \mathcal{H}}{\partial t} dt \tag{47}$$

From Lagrange's equations $\dot{\mathbf{p}} = \partial L / \partial \mathbf{q}$ and this gives us the Hamilton's equations since variations in all directions should be equal. □

In the case of mechanical equations we have $L = T - U$ and in the situation that the kinetic energy is in a quadratic form with respect to $\dot{\mathbf{q}}$ and we can see that the hamiltonian comes out to become $\mathcal{H} = T + U$ in this case.

Corollary 2.1.1. $d\mathcal{H}/dt = \partial \mathcal{H} / \partial t$, hence a system whose hamiltonian has no explicit time dependence, will have a law of conservation $\mathcal{H}(\mathbf{p}(t), \mathbf{q}(t)) =$ constant.

2.4 Liouville's Theorem and Poincare Recurrence

Definition 2.6 (Phase flow)**.** The phase flow is a one parameter group of transformations of a phase space

$$g^t : (\mathbf{p}(0), \mathbf{q}(0)) \to (\mathbf{p}(t), \mathbf{q}(t)) \tag{48}$$

Here $_{|}(t), \mathbf{q}(t)$ are solutions of Hamilton's equations.

Theorem 2.2 (Liouville's theorem)**.** Phase flow preserves volume of any region D in the phase space over which the transformations are applied. Let's say the region evolves with time as $D(t) = g^t D(0)$ and $v(t)$ is the volume of $D(t)$, this theorem states that $v(t)$ is constant in t.

As per the transformation law we get from equation 48 we would express the volume as a multiple variable integral. As shown below it takes the form of a determinant of a matrix. We can think of this making sense by seeing this as an oriented volume as such a thing holds for the volume formed by an object composed of the vectors originating from a point. Intuitively this can be built by the fact that exchanging two of the vectors orientation wise should make the volume flip sign and this fully antisymmetrized nature is the sum represented by the matrix determinant with each vector being a row.

$$v(t) = \int_{D(0)} \det \frac{\partial g^t \mathbf{x}}{\partial \mathbf{x}} dx \tag{49}$$

For small values of t we can express the following

$$\frac{\partial g^t \mathbf{x}}{\partial \mathbf{x}} = I + \frac{\mathbf{f}}{\partial \mathbf{x}} t + \mathcal{O}(t^2) \tag{50}$$

Lemma 2.3. For any matrix $A \in \mathbb{R}^{n\times n}$, we have the following for $t \to 0$

$$\det(I + At) = 1 + \text{tr}(A)t + \mathcal{O}(t^2) \tag{51}$$

The above lemma follows from the fact that the only terms first order in t can be shown to add up to the trace of A. The trace of $\partial \mathbf{f}/\partial \mathbf{x} = \nabla \cdot \mathbf{f}$ and so using this we can express volume at small t as follows

$$v(t) = \int_{D(0)} [1 + t\nabla \cdot \mathbf{f} + \mathcal{O}(t^2)] dx \tag{52}$$

From this we can now write the following theorem

Theorem 2.4. If $\nabla \cdot \mathbf{f} \equiv 0$, g^t is a volume preserving transformation.

Proof. From equation 52, it follows that $(dv/dt)|_{t=0} = \int_{D(o)} \nabla \cdot \mathbf{f} dx$. This implies that if the divergence of $\mathbf{f}$ vanishes the derivative of the volume w.r.t time vanishes at $t = 0$. Here $t = 0$ can be extended to any $t = t_0$ since there is nothing special in choosing 0 hence giving us that the volume is indeed preserved for such a transformation. □

Definition 2.7 (Liouville's formula)**.** For the linear system $\dot{\mathbf{x}} = A(t)\mathbf{x}$, the Wronskian determninant is given by $W = W_0 e^{\int \text{tr}}(A)dt$.

Proof. Taken from [15]. Let us define a matrix solution $\Phi : \mathbb{R} \to \mathbb{R}^{n\times n}$, which satisfies $\dot{\Phi}(t)A(t) = \Phi(t)$. Using the notation of determinant as a sum with the levi-civita tensor, it's easy to see that the derivative of $\det \Phi(t)$ would be the sum of the determinants with each row having been differentiated term-wise. Taking $A = \{a_{ij}\}$ and $\Phi = \{\phi_{ij}\}$ we can write the following

$$(\dot{\phi}_{i,1}, \ldots, \dot{\phi}_{i,n}) = \sum_{j=1}^{n} a_{i,j}(\phi_{i,1}, \ldots, \phi_{i,n})$$

In the determinant which replaces the ith row with it's derivative, we can subtract $\sum_{j=1,j\neq i}^{n} a_{i,j}(\phi_{j,1}, \ldots, \phi_{j,n})$ which is a linear combination of all other rows in the determinant and hence does not change the expression. This gives

$$\det \begin{pmatrix} \phi_{1,1} & \cdots & \phi_{1,n} \\ \vdots & \ddots & \vdots \\ \dot{\phi}_{i,1} & \cdots & \dot{\phi}_{i,n} \\ \vdots & \ddots & \vdots \\ \phi_{n,1} & \cdots & \phi_{n,n} \end{pmatrix} = \det \begin{pmatrix} \phi_{1,1} & \cdots & \phi_{1,n} \\ \vdots & \ddots & \vdots \\ a_{i,i}\phi_{i,1} & \cdots & a_{i,i}\phi_{i,n} \\ \vdots & \ddots & \vdots \\ \phi_{n,1} & \cdots & \phi_{n,n} \end{pmatrix} = a_{i,i} \det(\Phi)$$

Adding the above expression for all i we get the differential equation $\dot{W} = \text{tr}(A)W$ which solves to give the expected expression for initial state W_0. □

Theorem 2.5 (Poincaré's recurrence theorem)**.** Let g be a volume preserving transformation which is continuous and is a one-to-one mapping. Let D be some bounded region of euclidean space onto itself i.e., $gD = D$. Then in any neighbourhood U for any point in D, there exists a point $x \in U$ such that $\exists n > 0$ such that $g^n x \in U$

Proof. Let's pick some $U \subset D$ with $U \neq \emptyset$. Since g preserves volume, $\forall n, v(g^n U) = v(U)$ where $v(U)$ is the volume of the set U. All these images must lie in D since $gD = D$. Since we have an infinite number of such images, assuming none intersect each other this would require D to have infinite volume which is a contradiction.
Hence this requires that $\exists k, l \geq 0$ with $k > l$ such that $g^k U \cup g^l U \neq \emptyset$ which is the same as saying $g^{k-l} U \cup U \neq \emptyset$ hence proving the statement. □

Poincaré's statement can be strengthened to state that almost every moving point returns an infninte number of times to it's initial position (explicitly state in it's topological form [1]).
This theorem has some very interesting implications such as the claim that in a two chamber container with gas contained only in the first chamber, all the molecules would return back to their initial point in the first chamber after some time. The resolution of this seemingly paradoxical claim lies in the fact that this time scale may be well longer than the duration of the solar system's existence.
Some common examples for application of Poincaré's theorem are related to systems which involve an inherent periodicity such as movement on a circle or on a surface of a torus.
On a circle a transformation of an angle $2\pi\alpha$ if α is irrational is dense on the circle since $\forall \delta > 0, \exists n : |g^n x - x| < \delta$ due to the irrationality of α causing it never to be able to perform a proper loop.
Similarly for a system with two angles ϕ_1, ϕ_2 which are transforming as $g^t : (\phi_1, \phi_2) \to (\phi_1 + \alpha_1 t . \phi_2 + \alpha_2 t)$ one can fill up the phase space of the two variables which would be winding lines on a torus if α_1/α_2 is irrational. If their ratio is rational, there is some time where it would exactly return to the initial state but otherwise the winding lines would densely fill up the surface.

3 Manifolds

Figure 2: A nice and cozy manifold we happen to inhabit. Picture credits.

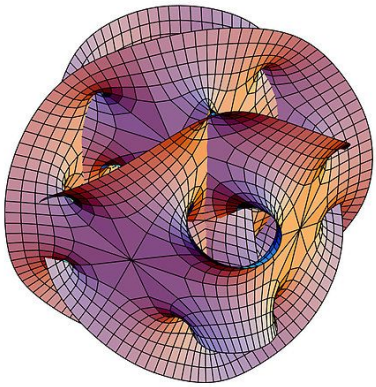

Figure 3: A complex four-space representation of the 'Calabi-Yau Manifold' by Stewart Dickson, which we also possibly inhabit. Picture credits.

3.1 What is a manifold?

Before we see what a manifold is, let's build toward some examples of similar such objects which are not. This section is in reference to [13], [14], [16], [10]. Let us first consider the 2 dimensional euclidean plane. Let's make a circle, a square and some n number of non intersecting circles on this flat plane. For whatever point we pick on the circle, we can take some open neighbourhood of points which also lie on that circle and map these to a normal open line segment. This map is onto and continuous and will exist for any such portion picked from the circle. Similarly this can be shown for the square too even around the sharp right angled corners since the orientation of the points in this manner does not matter and it can be parametrized and mapped. The n circles similarly can also be mapped since none intersect and hence every point would have some open neighbourhood which only has points from the same circle.
Things however get dicey if we try finding such mappings for two intersecting line segments. The point of intersection has no neighbourhood which has a continuous onto map to the open line segment. Since no matter how small the neighbourhood is, there will always be four directions of approach to the intersection, hence making continuous

onto mapping to a open line segment (which can only have two directions) impossible. This follows for any kind of intersection of objects that could have been mapped to the open line segment.
In 3 dimensions we have a similar mapping idea. Given a surface in 3D, we pick a point on it and want it's neighbourhood to have a continuous onto map to the open disc. We can see that analogous to the 2D case, a sphere, a torus, a cube and any n number of spheres will be able to satisfy this property. Again in the case of two spheres that touch, the point of contact will be problematic since no neighbourhood around it can be mapped as we want to an open disc. This mapping property is something which in a way is an important concept of what makes a manifold what it is.
We can now proceed to get into the actual definitions

Definition 3.1 (Manifold)**.** A n-manifold is a topological space in which each point has a neighbourhood which is homeomorphic (having a bijective continuous map) to the Euclidean space of dimension n.

This definition is essentially that of a *topological* manifold since this allows for the manifold to lack a certain structure which would be required for a differentiable manifold. To avoid certain pathalogical cases a topological manifold is also required to be hausdorff and second countable.

Definition 3.2 (Hausdorff space)**.** A Hausdorff space is a topological space where any two distinct points in this space there exists a neighbourhood for each point such that these neighbourhoods are disjoint.

Definition 3.3 (Second countable)**.** A topological space T is second-countable if there exists some countable collection $\mathcal{U} = \{U_i\}_{i=1}^{\infty}$ of open subsets of T such that any open subset of T can be written as a union of elements of some subfamily of $\mathcal{U}$.

Before we dive into differentiable manifolds we need to understand what charts are.

Definition 3.4 (Chart)**.** A chart is an open set U in $\mathbb{R}^n$ with an onto mapping $\phi : U \to \phi U \subset M$ where M is a n manifold. It is represented by this pair as (U, ϕ).

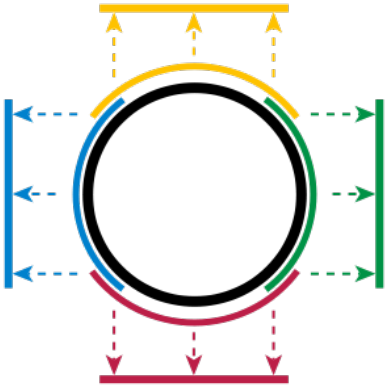

Figure 4: Mapping portions of a circle to open line segments. Picture taken from [16].

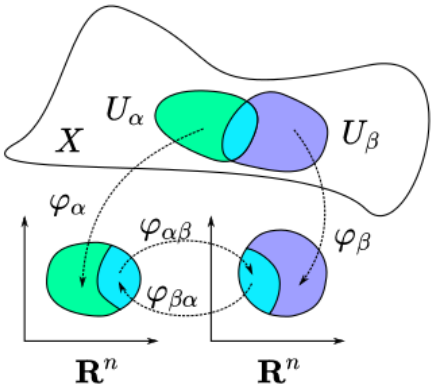

Figure 5: Compatible charts. Picture taken from [10]

Let's assume we have two points p and p' in charts U and U' respectively such that their image is the same point in M, then these points have neighbourhoods $V \subset U$, $V' \subset U'$ such that they have the same image in M. Hence this gives us a mapping $\phi'^{-1}\phi : V \to V'$ and the charts are called compatible if this function and it's inverse mapping are differentiable (see figure 5). An atlas is a union of compatible charts. Two atlases are referred to as equivalent if their union is also an atlas. A differentiable manifold is a class of equivalent atlases.
We can see that the euclidean space $\mathbb{R}^n$ is a manifold which has an atlas with only one chart. The n sphere given by $S^n = \{(x_1, \ldots, x_{n+1}) : \sum x_i^2 = 1\}$ has an atlas with two charts which are it's stereographic projections. One can see various examples of configuration spaces which come out to be manifolds such as a spherical pendulum is essentially having a configuration space of S^2. Similarly the configuration space of two pendulums which are both planar is the direct product $S^1 \times S^1 = T^2$ which is a torus.

Definition 3.5 (Embedded manifold)**.** We say M is an embedded k dimensional sub-manifold of $\mathbb{R}^n$ if in some neighbourhood U for each point $\mathbf{x} \in M$ there are $n - k$ functions $f_1 : U \to \mathbb{R}, \ldots f_{n-k} : U \to \mathbb{R}$ such that $f_i = 0$ and the gradients of all these functions are linearly independent at $\mathbf{x}$

It can be shown that every manifold is embedded in some euvlidean space. We can show that $SO(3)$, the special orthogonal group of 3×3 matrices are embedded in $\mathbb{R}^9$. To show this we first note that any member of $SO(3)$ satisfies

the equation

$$O \in SO(3) \implies OO^T = I_{3\times 3} \tag{53}$$

More importantly OO^T is a 3×3 symmetric matrix and so has six independent parameters and hence making it equal identity gives us six independent equations. This shows that the O matrix should have only 3 independent variables hence showing that it is 3 dimensional. However more importantly the six equations which constrain it are nothing but mappings from a neighbourhood of the point in the manifold to the set of real numbers and shows that we have $f_1 : U \to \mathbb{R} \dots f_6 : U \to \mathbb{R}$ and it's easy to see that they have linearly independent gradients at the solution hence showing that the 3 dimensional manifold $SO(3)$ is embedded in $\mathbb{R}^9$. It must be noted that the implies in equation 53 becomes an iff when the manifold is $O(3)$. The $SO(3)$ has an additional condition which makes it a transformation which preserves handedness and so is just the connected form of $O(3)$.

3.2 Tangent space and the tangent bundle

We have already seen what an embedded manifold is. Now let's say we have a k manifold embedded in $\mathbb{R}^n$. At any given point we have a k dimensional tangent space $TM_{\mathbf{x}}$ which is orthogonal to all the vectors ∇f_i where i goes from 1 to $n-k$. The vectors in this space are called tangent vectors to M at $\mathbf{x}$ and can also be defined as velocity vectors of curves in M

$$\dot{\mathbf{x}} = \lim_{t\to 0} \frac{\varphi(t) - \varphi(0)}{t}, \varphi : \mathbb{R} \to M, \varphi(0) = \mathbf{x} \tag{54}$$

Two curves $\varphi(t)$ and $\psi(t)$ are called equivalent if $\psi(0) = \varphi(0) = \mathbf{x}$ and $\lim_{t\to 0}(\varphi(t) - \psi(t))/t = 0$ in some chart. The tangent relationship will hold in all charts for the manifold since compatibility forces that they map to each other using a differentiable map and so the composition of the map on top of the curve would still obey the tangent relation.

Definition 3.6. A tangent vector to a manifold M at the point $\mathbf{x}$ is an equivalence class of curves $\varphi(t)$, with $\varphi(0) = \mathbf{x}$.

The tangent vector is essentially just $d\varphi/dt|_{t=0}$. Now one thing which happens here is that the union $\bigcup_{\mathbf{x}\in M} TM_{\mathbf{x}}$ has a natural differential manifold structure which has twice the dimension of M. This new manifold has twice as many dimensions since any point on it is defined by a k dimensional tangent vector $\dot{\mathbf{x}}$ on a k dimensional point $\mathbf{x}$ in M.
This manifold TM is referred to as the tangent bundle. The mapping $p : TM \to M$ takes the tangent vector to the point in M where it is tangent to it and is called the natural projection. The inverse image of a point in M is it's tangent space $TM_{\mathbf{x}}$ and this space is called the fiber of the tangent bundle over that point.

Definition 3.7 (Reimannian manifold)**.** A differentiable manifold with a fixed positive definite quadratic form $\langle \zeta, \zeta \rangle$ on every tangent space $TM_{\mathbf{x}}$ is called as a Reimannian manifold and the quadratic form is referred to as the Reimannian metric.

The concept of the Reimannian metric is extensively used for manifolds in general relativity where one writes $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$.
One can create say some mapping $f : M \to N$ where M and N are manifolds and in local coordinates of M and N we have f to be differentiable. We can proceed to define a new map between the tangent bundles of these manifolds.

Definition 3.8. The derivative of a differentiable mapping $f : M \to B$ at some point $\mathbf{x} \in M$ is the linear map

$$f_{*\mathbf{x}} : TM_{\mathbf{x}} \to TN_{f(\mathbf{x})} \tag{55}$$

If we take some $v \in TM_{\mathbf{x}}$ and $\varphi : \mathbb{R} \to M$ with $\varphi(0) = \mathbf{x}$ with the velocity vector for this curve being v, we would define the velocity vector of the curve $f \circ \varphi : \mathbb{R} \to N$ as

$$f_{*\mathbf{x}} v = \left. \frac{df(\varphi(t))}{dt} \right|_{t=0} \tag{56}$$

We can check that the value of the RHS has no dependence on φ since on applying chain rule and the constraints that $d\varphi/dt|_{t=0} = v$ and $\varphi(0) = \mathbf{x}$ we get $f_{*\mathbf{x}} v = \nabla_{\mathbf{x}} f v$.
We could take the union of all these mappings $F_{*\mathbf{x}}$ and just denote $F_* TM \to TN$ by keeping $F_* v = f_{*\mathbf{x}} v$ for $v \in TM_{\mathbf{x}}$. We can see that this is well defined as each point in TM would have a point in M along with a vector in $TM_{\mathbf{x}}$ associated with it. Hence all this map does is take a point $(\mathbf{x}, \dot{\mathbf{x}})$ to $(f(\mathbf{x}), \nabla_{\mathbf{x}} f \dot{\mathbf{x}})$ and so is clearly differentiable.

3.3 Making physical sense of manifolds

While we were building the tangent space and tangent bundle, we actually were getting to the very familiar Lagrangian dynamics which actually is defined on a tangent bundle. If M is a differentiable mnanifold and TM is the tangent

bundle, $L : TM \to \mathbb{R}$ is the lagrangian (assuming it has no time dependence). The extremal defined for the functional $\Phi(\gamma) = \int_{t_0}^{t_1} L(\dot{\gamma})dt$ where $\dot{\gamma}(t) \in TM_{\gamma(t)}$.
The evolution of these coordinates is defined by the Euler-Lagrange equations in equation 39.

Definition 3.9. A lagrangian system on a Reimannian manifold is called *natural* $L = T_U$

One can now have certain holonomic constraints in the system. Since every manifold can be embedded in a certain Euclidean space and one can define the potential as something of the form $U' = U + N\mathbf{q}_2^2$ where the coordinate $\mathbf{q}_2$ is chosen in a way that it is non zero only outside the manifold hence taking $N \to \infty$ the potential rapidly grows outside of M.
Now we must note that this is all for only autonomous lagrangian systems since L has no explicit time dependence.

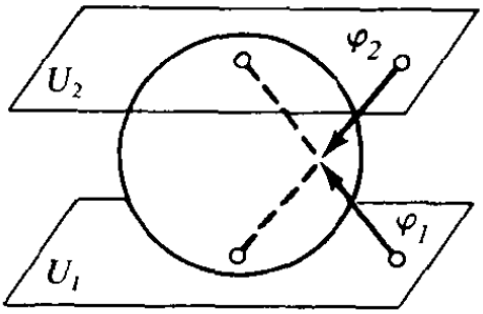

Figure 6: Stereographic projections of a sphere. Picture taken from [14]

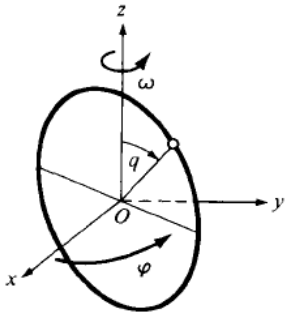

Figure 7: A bead constrained to move on a rotating circle. Picture taken from [14]

More generally we would denote $L : TM \times \mathbb{R} \to \mathbb{R}$. The time dependent manifold is mapped to the euclidean space using $i : M \times \mathbb{R} \to E^{3n}$ for a system of n particles. One can have situations where constraints itself are time dependent such as a bead on a rotating circle. For this case $i(q,t) = (r\sin(q)\cos(\omega t), r\sin(q)\sin(\omega t), r\cos(q))$ and so it effectively is just a one dimensional problem now and the kinetic energy from the rotating circle can be grouped into the potential to make it $V = mgr\cos(q) - m\omega^2 r^2 \sin^2(q)/2$.

3.4 Some additional notes

So one may have noted that the topological manifold which is supposed to be the most general definition for any manifold. This definition however included two important points to eliminate pathalogical cases which were being Hausdorff and second countable. As we saw in section 1.1, closeness can be claimed without concerning oneself with a distance metric for that space.
There are some interestic non-Hausdorff spaces which show why it's best to have this as an additional condition. The most common example is that of a real line with two origins. It is defined with matching the real line at all points except at the origin which is left to two distinct points. One cannot picture it physically but the broken definition makes the two points for origin indistinguishable since their neighbourhoods will always have some non empty intersection. This line is represented by the quotient space of $\mathbb{R} \times \{a\}$ and $\mathbb{R} \times \{b\}$ with the equivalence relation $(x,a) \sim (x,b)$, $x \neq 0$ and we can see that 0_a has no disjointed neighbourhood from 0_b. More of such examples are termed as non-hausdorff manifolds and so here are our pathological examples.

4 Symmetry

We see symmetry arise in lots of situations in nature such as rotational invariance for wavefunctions when they have a potential only depending on radius. One thing to note is that every symmetry is associated with some conserved quantity and interestingly symmetry breaking can cause some interesting things too.

4.1 Noether's theorem

Definition 4.1. A lagrangian L on a manifold M admits the mapping $h : M \to M$ if $\forall v \in TM$, $L(h_* v) = L(v)$

For example the transformation $h : (x_1, x_2, x_3) \to (x_1 + s, x_2, x_3)$ on $L = (m/2)(\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2) - U(x_2, x_3)$ does not change it and hence it admits h. Using this we state Noether's theorem

Theorem 4.1 (Noether's theorem)**.** If a lagrangian L on M admits the on-parameter group of diffeomorphisms $h^s : M \to M$, $s \in \mathbb{R}$ then the lagrangian system of equations corresponding to L has a first integral $I : TM \to \mathbb{R}$ which in local coordinates $\mathbf{q}$ on M can be written as

$$I(\mathbf{q}, \dot{\mathbf{q}}) = \frac{\partial L}{\partial \dot{\mathbf{q}}} \frac{dh^s(q)}{ds}\bigg|_{s=0} \tag{57}$$

Proof. Let $\varphi : \mathbb{R} \to M$, $\mathbf{q} = \varphi(t)$ be a solutions to the Lagrange's equations. Since L admits h^s, this means $h^s\varphi : \mathbb{R} \to M$ also satisfies the Lagrange's equations for all $s \in \mathbb{R}$. Since s is a continuous variable, this means that L does not depend explicitly on s hence we can write the following for $\mathbf{q} = h^s(\varphi(t))$

$$\frac{\partial L(\Phi, \dot{\Phi})}{\partial s} = \frac{\partial L}{\partial \mathbf{q}} \frac{\partial \mathbf{q}}{\partial s} + \frac{\partial L}{\partial \dot{\mathbf{q}}} \frac{\partial \dot{\mathbf{q}}}{\partial s} = 0 \tag{58}$$

However it is known that $\mathbf{q}$ is a solution to the Lagrange equations and hence we can write $\dfrac{\partial L}{\partial \mathbf{q}} = \dfrac{d}{dt}\dfrac{\partial L}{\partial \dot{\mathbf{q}}}$ and so we would write the following from the previous equation

$$\left(\frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{q}}}\right)\frac{\partial \mathbf{q}}{\partial s} + \frac{\partial L}{\partial \dot{\mathbf{q}}}\frac{\partial \dot{\mathbf{q}}}{\partial s} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}}\frac{\partial \mathbf{q}}{\partial s}\right) = 0 \tag{59}$$

Hence giving us the first integral to be the claimed expression. □

Now let us first understand what exactly a first integral by referring to it's definition [11].

Definition 4.2 (First integral)**.** Given a system of differential equations $\Phi(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) = 0$, the first integral is a non constant function $I : TM \to \mathbb{R}$ such that $dI(\mathbf{x}(t), \dot{\mathbf{x}}(t))/dt = 0$ for any solution of $\Phi(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) = 0$.

The reason it is referred to as an integral actually is due to the fact that in earlier times mathematicians used to integrate ODEs to get solutions and so the term "integral" is associated with solution in this sense. While one may see this quantity as just the conserved quantity due to symmetry (which it infact really is), there is more interesting physical observations to it. For autonomous systems, one can think of the first integral to represent a scalar potential and solutions represent the equipotential surfaces.

If a system admits rotations around an axis say e_1, the angular momentum with respect to that axis is conserved. We can see this by observing that if h^s is the rotation, $(dh^s\mathbf{x}_i/ds)|_{s=0} = e_1 \times \mathbf{x}_i$ and so from this it follows that $I = \sum_i \partial L/\partial \dot{\mathbf{x}}_i \cdot (e_1 \times \mathbf{x}_i)$ which is simply the total angular momentum about that axis since $\partial L/\partial \dot{\mathbf{x}}_i = m_i \dot{\mathbf{x}}_i$ (assuming the system is natural).

While our proof dealt with an autonomous system, it can be extended to non autonomous systems too. If we define a new manifold $M_1 = M \times \mathbb{R}$, we can also define a new lagrangian for this system as $L_1 : TM_1 \to \mathbb{R}$ by $L_1 = L\dfrac{dt}{d\tau}$ where τ is the parameter which we are using to parametrize everything including the actual time t. Hence we would write

$$L_1\left(\mathbf{q}, t, \frac{d\mathbf{q}}{d\tau}, \frac{dt}{d\tau}\right) = L\left(\mathbf{q}, \frac{d\mathbf{q}}{d\tau}, t\right)\frac{dt}{d\tau} \tag{60}$$

We can now apply Noether's on L_1 to get the first integral $I_1 : TM_1 \to \mathbb{R}$ with no issues and since $\int L dt = \int L_1 d\tau$ this integral reduces to $I : TM \times \mathbb{R} \to \mathbb{R}$. A particular interesting case here is when L has no explicit time dependence in which case translation of the time coordinate is a transformation that does not change the lagrangian and the quantity that is conserved here is the energy (a very common case in dynamical systems).

4.2 Stress-energy tensor

So far we have abstained from using the field theory notation with tensors and indices but now we will introduce the stress energy tensor (or energy momentum tensor) from the lagrangian density motivation. Previously we used continuous transformations on tangent vectors on a manifold, we will now instead deal with continuous transformations on fields $\phi(x)$. Let's write any continuous infinitesimal transformation as $\phi'(x) = \phi(x) + \alpha\Delta\phi(x)$ where α is an infinitesimal parameter and $\Delta\phi$ is a deformation of the field. The lagrangian would have to transform invariant upto a 4-divergence to be conserved as we would have $\mathcal{L}(x) = \mathcal{L}(x) + \alpha\partial_\mu \mathcal{J}^\mu(x)$. Similar to how we proved Noether's theorem, we use the Euler Lagrange equations and we get the following current

$$j^\mu(x) = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\Delta\phi - \mathcal{J}^\mu \tag{61}$$

This current follows the conservation law $\partial_\mu j^\mu = 0$ which implies that there is some kind of conserved charge which can be expressed as $Q = \int j^0 d^3x$. In case the Lagrangian is invariant to the transformation we have $\mathcal{J}^\mu = 0$ and we can see that j^μ is nothing but the first integral in equation 57.
Now let's say that we are transforming the coordinates by some translation which effectively changes the field too. We would represent the transformation as follows

$$\phi(x+a) = \phi(x) + a^\mu \partial_\mu \phi(x) \tag{62}$$

Given that the lagrangian is scalar we would have it transform as $\mathcal{L}' = \mathcal{L} + a^\nu \partial_\mu(\delta^\mu_\nu \mathcal{L})$. On comparing with the previous case this is essentially a non zero $\mathcal{J}^\mu$ and to take this into account one can obtain four separately conserved currents by the following tensor

$$T^\mu_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial_\nu \phi - \mathcal{L}\delta^\mu_\nu \tag{63}$$

This gives us the conserved charge to be the Hamiltonian which is conserved with time transalations.

$$H = \int T^{00} d^3x = \int \mathcal{H} d^3x \tag{64}$$

The other spatial transalations are associated iwth a physical momentum carried by the field

$$P^i = \int T^{0i} d^3x = -\int \frac{\partial \mathcal{L}}{\partial \dot{\phi}(x)} \partial_i \phi d^3x \tag{65}$$

4.3 Symmetry breaking

A symmetric world would be dull [17]

Let us consider the following lagrangian density

$$\mathcal{L} = \frac{1}{2}\left[(\partial]\vec{\varphi})^2 - \mu^2\vec{\varphi}^2\right] - \frac{\lambda}{4}(\vec{\varphi}^2)^2 \tag{66}$$

Here $\vec{\varphi}$ is an N component vector. We can see that L admits all transformations which rotate $\vec{\varphi}$ which is essentially $O(N)$ symmetry. Here $\vec{\varphi}$ is essentially a field. One can manually remove the symmetry present by adding terms which simply depend on certain components alone however there is a way for spontaneous symmetry breaking if we define our lagrangian slightly differently.

$$\mathcal{L}' = \frac{1}{2}\left[(\partial]\vec{\varphi})^2 + \mu^2\vec{\varphi}^2\right] - \frac{\lambda}{4}(\vec{\varphi}^2)^2 \tag{67}$$

This is essentially what looks like a double well potential's lagrangian where we have

$$V(\varphi) = -\frac{1}{2}\mu^2\vec{\varphi}^2 + \frac{\lambda}{4}(\vec{\varphi}^2)^2 \tag{68}$$

This has two minimas which are symmetric about the origin in the one dimensional case. In quantum mechanics, presence in one of the minimas would imply the possibility of being able to tunnel to the other minima too. However the tunneling barrier is now $[V(0) - V(\pm v)]\int d^Dx$ where D is the dimensions of space (1 in this case) and so tunneling is shut down as this barrier is effectively infinite. Hence the ground state must make a choice between two equivalent situations showing that naturally the symmetry is destroyed. If we choose the ground state to be at $+v$ we can write $\varphi = v + \varphi'$ and expanding about φ' we get

$$\mathcal{L}' = \frac{\mu^4}{4\lambda} + \frac{1}{2}(\partial\varphi')^2 - \mu^2\varphi'^2 - \mathcal{O}(\varphi'^3) \tag{69}$$

The shifted field φ' has a mass of $\sqrt{2}\mu$ on comparing with equation 68 which is an interesting point to note. On extending to two dimensions, we get the sombrero potential (figure 8) which has $O(2)$ symmetry that will be broken for the ground state since one would have choose an angle and so this is a breaking of continuous symmetry. Since the choice is arbitrary we can pick a point with $\varphi_1 = v$ and $\varphi_2 = 0$ and on expanding we get

$$L' = \frac{\mu^4}{4\lambda} + \frac{1}{2}(\partial\vec{\varphi'})^2 - \mu^2\varphi'^2_1 - \mathcal{O}(\varphi'^3) \tag{70}$$

We can see now that the field φ_2 seems to have no mass and $\varphi = \sqrt{2}\mu$.

Theorem 4.2 (Goldstone's theorem)**.** Whenever a continuous symmetry is spontaneously broken, massless fields, known as Nambu-Goldstone bosons emerge.

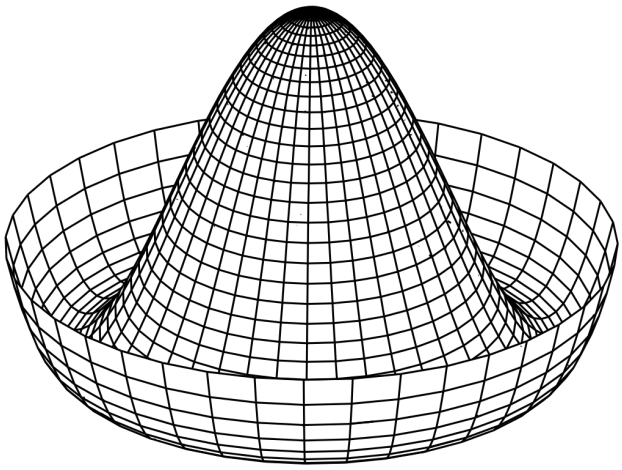

Figure 8: The sombrero potential. Picture credits.

.

Proof. Every continuous symmetry has a conserved charge Q which generates a symmetry which relates to the Hamiltonian as $[H, Q] = 0$. Let's say the ground state is $|0\rangle$ and the Hamiltonian is shifted by a constant so that $H|0\rangle = 0$. Normally the ground state would have to be invariant under symmetry transform however that does not hold here so $Q|0\rangle \neq |0\rangle$. So let's say we want the energy of this transformed state which would be $HQ|0\rangle = (HQ - QH + QH)|0\rangle = 0$. More importantly from the Fabri-Picasso theorem, $Q|0\rangle = 0$ as otherwise it would not be able to exist in the Hilbert space. This is due to $Q(t) = \int dx^D J^\mu(t)$ and the correlation function $\langle 0|QQ|0\rangle$ will end up going to infinity unless it's zero itself. □

5 Klein-Gordan Field

Before we go on to the actual nature of this field, it is important to understand what makes this necessary in the first place. It is seen that when writing a single particle relativistic wave equation (such as the Klein-Gordan equation or the Dirac equation) there are negative energy states which arise. These are resolved in the Dirac equation using a concept called the Dirac Sea [5].
The necessity of a multiparticle theory arises from having causality. Taking the propagation amplitude of a free particle from one point in space $\mathbf{x}_0$ to another $\mathbf{x}$. We would have the following expression for the same

$$U(t) = \langle \mathbf{x}| e^{-\iota Ht} |\mathbf{x}_0\rangle \tag{71}$$

One can expand this in the relativistic expression of $E = \sqrt{p^2 + m^2}$ and on reduction of the integral the following expression is obtained

$$U(t) = \frac{1}{2\pi^2|\mathbf{x} - \mathbf{x}_0|} \int_0^\infty dp \sin(p|\mathbf{x} - \mathbf{x}_0|) p e^{-\iota t\sqrt{p^2+m^2}} \tag{72}$$

This can be evaluated in terms of Bessel functions [8]. What concerns us is the asymptotic behaviour well outside the light cone $\mathbf{x}^2 >> t^2$. The phase function would have a stationary point at $p = \iota mx/\sqrt{x^2 - t^2}$ and so the contour integral can be pushed to this point and we get that $U(t) \sim e^{-m\sqrt{x^2-t^2}}$ upto some rational function of x and t. This shows that there is a propagation amplitude that is vanishing but non zero which goes outside the light cone and that is not a good thing clearly as we no longer have causality.
The Klein Gordan equation is the equations of motion of a very basic lagrangian density shown below

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2}m^2\phi^2 \tag{73}$$

Which gives us the Klein-Gordan equation below

$$(\partial^\mu \partial_\mu + m^2)\phi = 0 \tag{74}$$

5.1 Klein Gordan field as Harmonic Oscillators

Here we will quantize classical field theory by promoting ϕ and $\pi = \partial\mathcal{L}/\partial\dot{\phi}$ (canonical momenta) into operators much like how we treat q (position) and p (momentum) as operators in quantum mechanics. Similar to our very familiar commutation relations in equation 14 we write the following

$$\begin{aligned}[\phi(\mathbf{x}), \pi(\mathbf{y})] &= \iota\delta^{(3)}(\mathbf{x}-\mathbf{y}) \\ [\phi(\mathbf{x}), \phi(\mathbf{y})] &= [\pi(\mathbf{x}), \pi(\mathbf{y})] = 0\end{aligned} \tag{75}$$

Now our aim would be to construct the Hamiltonian that describes the system for the Klein-Gordan field. We can start that off by taking the Fourier transform of the field

$$\phi(\mathbf{x}, t) = \int \frac{d^3p}{(2\pi)^3} e^{\iota\mathbf{p}\cdot\mathbf{x}} \phi(\mathbf{p}, t) \tag{76}$$

The Klein-Gordan equation in equation 74 becomes into

$$\left[\frac{\partial^2}{\partial t^2} + (|\mathbf{p}|^2 + m^2)\right] \phi(\mathbf{p}, t) = 0 \tag{77}$$

However this really is the same as the equation of a harmonic oscillator with frequency of $\omega_{\mathbf{p}} = \sqrt{|\mathbf{p}|^2 + m^2}$. Now similar to how we have written the ladder operators for a simple harmonic oscillator, we can do the same for every individual Fourier mode of the field hence treating each mode as an independent oscillator with it's own ladder operations.

$$\phi(\mathbf{x}) = \int \frac{d^3}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \left(a_{\mathbf{p}} e^{\iota\mathbf{p}\cdot\mathbf{x}} + a^\dagger_{\mathbf{p}} e^{-\iota\mathbf{p}\cdot\mathbf{x}}\right) \tag{78}$$

$$\pi(\mathbf{x}) = \int \frac{d^3}{(2\pi)^3} (-\iota) \sqrt{\frac{\omega_{\mathbf{p}}}{2}} \left(a_{\mathbf{p}} e^{\iota\mathbf{p}\cdot\mathbf{x}} - a^\dagger_{\mathbf{p}} e^{-\iota\mathbf{p}\cdot\mathbf{x}}\right) \tag{79}$$

We will find it useful to write $a^\dagger_{\mathbf{p}} e^{-\iota\mathbf{p}\cdot\mathbf{x}} = a^\dagger_{-\mathbf{p}} e^{\iota\mathbf{p}\cdot\mathbf{x}}$ which would make further calculations simpler. We note that we have the following commutation relation hold

$$[a_{\mathbf{p}}, a^\dagger_{\mathbf{p}'}] = (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{p}') \tag{80}$$

It can be checked using this commutator relation that it equivalently gives $[\phi(\mathbf{x}), \pi(\mathbf{x}')] = \iota\delta^{(3)}(\mathbf{x} - \mathbf{x}')$. Given a ϕ and π we can construct the Hamiltonian from these as follows

$$H = \int d^3x \mathcal{H} = \int d^3x \left[\frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\phi)^2 + \frac{1}{2}m^2\phi^2\right] \tag{81}$$

Using this we can now finally construct the Hamiltonian for the Klein-Gordan Field.

$$\begin{aligned}H &= \int d^3x \int \frac{d^3p d^3p'}{(2\pi)^6} e^{\iota(\mathbf{p}+\mathbf{p}')\cdot\mathbf{x}} \left\{-\frac{\sqrt{\omega_{\mathbf{p}}\omega_{\mathbf{p}'}}}{4}(a_{\mathbf{p}} - a^\dagger_{-\mathbf{p}})(a_{\mathbf{p}} - a^\dagger_{-\mathbf{p}}) + \frac{m^2 - \mathbf{p}\cdot\mathbf{p}'}{4\sqrt{\omega_{\mathbf{p}}\omega_{\mathbf{p}'}}}(a_{\mathbf{p}} + a^\dagger_{-\mathbf{p}})(a_{\mathbf{p}} + a^\dagger_{-\mathbf{p}})\right\} \\ &= \int \frac{d^3p}{(2\pi)^3} \omega_{\mathbf{p}} \left(a^\dagger_{\mathbf{p}} a_{\mathbf{p}} + \frac{1}{2}[a_{\mathbf{p}}, a^\dagger_{\mathbf{p}}]\right)\end{aligned} \tag{82}$$

One may notice that the second term is proportional to $\delta(0)$ which is essentially infinite which while disturbing is not a real issue since energy differences are what matter. This is due to the sum of all the zero point energies being summed over for all the Harmonic oscillators. However one cannot ignore this zero point energy always as one can see in the Casimir effect [4].

$$[H, a^\dagger_{\mathbf{p}}] = \omega_{\mathbf{p}} a^\dagger_{\mathbf{p}} \qquad [H, a_{\mathbf{p}}] = \omega_{\mathbf{p}} a^\dagger_{\mathbf{p}} \tag{83}$$

We now see that we have the state $|0\rangle$ which we will treat as the ground state for all $\mathbf{p}$ since $a_{\mathbf{p}}|0\rangle = 0$ and will be referred to as the vacuum. We can apply creation operators on this to create states $a^\dagger_{\mathbf{p}} a^\dagger_{\mathbf{q}} \ldots |0\rangle$ which would have energy $\omega_{\mathbf{p}} + \omega_{\mathbf{q}} + \ldots$ and these are the states that are present in the spectrum. We can on this basis create a total momentum operator as $\int \frac{d^3p}{(2\pi)^3\mathbf{p}} a^\dagger_{\mathbf{p}} a_{\mathbf{p}}$. We can call these excitations created by the creation operations to simply be particles each with a different momentum wavefunction. The states $a_{\mathbf{p}} a_{\mathbf{q}}|0\rangle$ and $a_{\mathbf{q}} a_{\mathbf{p}}|0\rangle$ are equivalent since these operations commute and also a single mode $\mathbf{p}$ can contain arbitrarily many particles which would be driven by applying any number of $a^\dagger_{\mathbf{p}}$ and so these particles obey Bose-Einstein statistics.

We would define the vaccum state to be normalized hence $\langle 0\rangle\, 0 = 1$. However if we similarly define our states as $|\mathbf{p}\rangle = a^\dagger_{\mathbf{p}} |0\rangle$ upto a constant and $\langle \mathbf{p}\rangle\, \mathbf{q} = (2\pi)^3\delta^(3)(\mathbf{p}-\mathbf{q})$ the normalization does not remain Lorentz invariant. This is due to the fact that $\delta(f(x) - f(x_0)) = \frac{1}{|f'(x_0)|}\delta(x - x_0)$ and so to correct for this we simply define the states as

$$|\mathbf{p}\rangle = \sqrt{2E_{\mathbf{p}}}a^\dagger_{\mathbf{p}} |0\rangle \qquad\qquad \langle \mathbf{p}\rangle\, \mathbf{q} = 2E_{\mathbf{p}}(2\pi)^3\delta^(3)(\mathbf{p}-\mathbf{q}) \tag{84}$$

Since we have $E_{\mathbf{p}}\delta^(3)(\mathbf{p}-\mathbf{q})$ as lorentz invariant. However introducing this extra factor will also require we define the completeness relation for one particle as

$$\mathbb{I}_{1\text{ particle}} = \int \frac{d^3p}{(2\pi)^3} |\mathbf{p}\rangle \frac{1}{2E_{\mathbf{p}}} \langle \mathbf{p}| \tag{85}$$

Infact the integral $\int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}}$ is a Lorentz invariant integral in the sense that if $f(p)$ is Lorentz invariant, so is $\int \frac{d^3p}{(2\pi)^3}\frac{f(p)}{2E_{\mathbf{p}}}$. This can be seen as the integral over the $p^0 > 0$ branch of the hyperboloid $p^2 = m^2$ in the 4-momentum space if written as follows

$$\int \frac{d^3p}{(2\pi)^3}\frac{f(p)}{2E_{\mathbf{p}}} = \int \frac{d^4p}{(2\pi)^3}\frac{f(p)}{\delta}(p^2 - m^2)\big|_{p^0>0} \tag{86}$$

This finally leads us to interpretation of $\phi(x)|0\rangle$. As per how ϕ is written in equation 80 we get

$$\phi(\mathbf{x})|0\rangle = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}}e^{-\iota\mathbf{p}\cdot\mathbf{x}}|\mathbf{p}\rangle \tag{87}$$

We will be treating this as $|\mathbf{x}\rangle$ since we can see for the non relativistic case $E_{\mathbf{p}}$ is nearly constant hence this simply sums all plane waves hence would give a point particle.

5.2 Solving the causality problem

A thing to note is that the previous subsection stuck to the Schrödinger picture. We will now proceed to use the Heisenberg picture. Using the Heisenberg equations of motion we can write the following

$$\begin{aligned}
\iota\frac{\partial}{\partial t}\phi(\mathbf{x},t) &= [\phi(\mathbf{x},t), \int d^3x' \left\{\frac{1}{2}\pi^2(\mathbf{x}',t) + \frac{1}{2}(\nabla\phi(\mathbf{x}',t))^2 + \frac{1}{2}m^2\phi^2(\mathbf{x}',t)\right\}] \\
&= \int d^3x' \left(\iota\delta^(3)(\mathbf{x}-\mathbf{x}')\pi(\mathbf{x}',t)\right) \\
&= \iota\pi(\mathbf{x},t)
\end{aligned} \tag{88}$$

$$\begin{aligned}
\iota\frac{\partial}{\partial t}\pi(\mathbf{x},t) &= [\pi(\mathbf{x},t), \int d^3x' \left\{\frac{1}{2}\pi^2(\mathbf{x}',t) + \frac{1}{2}\phi(\mathbf{x}',t)(-\nabla^2 + m^2)\phi(\mathbf{x}',t)\right\}] \\
&= \int d^3x' \left(\iota\delta^(3)(\mathbf{x}-\mathbf{x}')(-\nabla^2 + m^2)\phi(\mathbf{x}',t)\right) \\
&= -\iota(-\nabla^2 + m^2)\phi(\mathbf{x},t)
\end{aligned} \tag{89}$$

Using the commutation relation of H and $a_{\mathbf{p}}$ we can write $Ha_{\mathbf{p}} = a_{\mathbf{p}}(H - E_{\mathbf{p}})$ which can be extended to $H^n a_{\mathbf{p}} = a_{\mathbf{p}}(H - E_{\mathbf{p}})^n$. Using this we get the identities

$$e^{\iota Ht}a_{\mathbf{p}}e^{-\iota Ht} = a_{\mathbf{p}}e^{-\iota E_{\mathbf{p}}t} \qquad\qquad e^{\iota Ht}a^\dagger_{\mathbf{p}}e^{-\iota Ht} = a^\dagger_{\mathbf{p}}e^{\iota E_{\mathbf{p}}t} \tag{90}$$

We would similarly get the same identities using the total momentum operator $\mathbf{P}$ instead of H and $\mathbf{p}$ instead of $E_{\mathbf{p}}$. This lets us relate $phi(x)$ to $\phi(0)$ as

$$\begin{aligned}
\phi(x) &= e^{\iota(Ht - \mathbf{P}\cdot\mathbf{x})}\phi(0)e^{-\iota(Ht-\mathbf{P}\cdot\mathbf{x})} \\
&= e^{\iota P\cdot x}\phi(0)e^{-\iota P\cdot x}
\end{aligned} \tag{91}$$

Here the notation used is $P^\mu = (H, \mathbf{P})$. We must note that both $e^{-\iota p^0 t}$ and $e^{\iota p^0 t}$ appear in the expressions which shows positive and negative energy modes respectively. We will refer to these as positive and negative frequency modes. The positive frequency solution has it's coefficient as the operator which destroys a particle in that single

particle wavefunction and the negative mode has one which creates a particle.
As per how the operators have been defined, the amplitude for a particle to propogate to y from x is $\langle 0|\,\phi(x)\phi(y)\,|0\rangle$ which we will call as $D(x-y)$. Since each operator ϕ is composed of the ladder operators, only the term $\langle 0|\,a_{\mathbf{p}}a^{\dagger}_{\mathbf{q}}\,|0\rangle = (2\pi)^3\delta^3(\mathbf{p}-\mathbf{q})$ will be non zero. The expression which is left is

$$D(x-y) = \langle 0|\,\phi(x)\phi(y)\,|0\rangle = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}}e^{-\iota p\cdot(x-y)} \tag{92}$$

We already know that this kind of integral is Lorentz invariant. Let's say that $x-y$ is timelike separated we get $D(x-y)\sim e^{-\iota mt}$ taking $t\to\infty$. When they are spatially separated ($\mathbf{x}-\mathbf{y}=\mathbf{r}$) the integral becomes as follows

$$\begin{aligned} D(x-y) &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}}e^{\iota\mathbf{p}\cdot\mathbf{r}} \\ &= \frac{2\pi}{(2\pi)^3}\int_0^{\infty} dp\frac{p^2}{2E_{\mathbf{p}}}\frac{e^{\iota pr}-e^{-\iota pr}}{\iota pr} \\ &= \frac{-\iota}{2(2\pi)^2 r}\int_{-\infty}^{\infty} dp\frac{pe^{\iota pr}}{\sqrt{p^2+m^2}} \end{aligned} \tag{93}$$

There are two poles here which are $p=\pm\iota m$ so we can push the contours around one of the poles and substitute $\rho=-\iota p$ to get $D(x-y)\sim e^{-mr}$ when $r\to\infty$. Seeing these expressions it seems that our aim to solve the causality problem which we made so apparent does not really happen. However in the context of quantum mechanics, one would have to refer to causality as the act of one measurement affecting another. One knows that measurements that have a zero commutator do not affect each other and so we can pick the simplest measurement here which is ϕ and find the value of $[\phi(x),\phi(y)]$

$$\begin{aligned} [\phi(x),\phi(y)] &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{p}}}}\int \frac{d^3q}{(2\pi)^3}\frac{1}{\sqrt{2E_{\mathbf{q}}}}\left[(a_{\mathbf{p}}e^{-\iota p\cdot x}+a^{\dagger}_{\mathbf{p}}e^{\iota p\cdot x}),(a_{\mathbf{q}}e^{-\iota q\cdot y}+a^{\dagger}_{\mathbf{q}}e^{\iota q\cdot y})\right] \\ &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\mathbf{p}}}(e^{-\iota p\cdot(x-y)}-e^{\iota p\cdot(x-y)}) \\ &= D(x-y)-D(y-x) \end{aligned} \tag{94}$$

Now if $(x-y)^2<0$, we can perform a Lorentz transformation on the second term (the integral is Lorentz invariant so this shouldn't change anything) which takes $(x-y)$ to $(y-x)$. Such a transformation is only possible if they have a spacelike separation hence showing that in such a case the commutator is zero so causality is held.
Also we additionally define the green's function of the Klein-Gordan operator to be $D_R(x-y)=\theta(x^0-y^0)\,\langle 0|\,[\phi(x),\phi(y)]\,|0\rangle$. We can see that

$$\begin{aligned} (\partial^2+m^2)D_R(x-y) &= (\partial^2\theta(x^0-y^0))\,\langle 0|\,[\phi(x),\phi(y)]\,|0\rangle + 2(\partial_\mu\theta(x^0-y^0))(\delta^\mu\,\langle 0|\,[\phi(x),\phi(y)]\,|0\rangle) \\ &= -\delta(x^0-y^0)\,\langle 0|\,[\pi(x),\phi(y))]\,|0\rangle + 2\delta(x^0-y^0)\,\langle 0|\,[\pi(x),\phi(y))]\,|0\rangle \\ &= -\iota\delta^{(}4)(x-y) \end{aligned} \tag{95}$$

On noting the exact form of D_R we find that a different pole prescription called the Feynman prescription makes it far more convenient to write which is

$$D_F(x-y) = \int \frac{d^4p}{(2\pi)^4}\frac{\iota}{p^2-m^2+\iota\varepsilon}e^{-\iota p\cdot(x-y)} \tag{96}$$

Here ϵ is just an infinitesimal number and while icky to define this way, we treat any constant multiplied to it as just ϵ. The poles would now be at $p^0=\pm(E_{\mathbf{p}}-\iota\epsilon)$. When $x^0>y^0$ we can close the contour from below and vice versa for $x^0<y^0$ which makes $D_F(x-y)=D((x-y)\mathrm{sgn}(x^0-y^0))$ and so has the time ordering inbuilt into it.

5.3 Complex Klein-Gordan Field

The case of a complex Klein Gordan field happens to show some interesting properties. When the complex scalar field is quantized, $\phi(x)$ creates positively charged particles and destroys negatively charged particles while $\phi^{\dagger}(x)$ does the opposite. As a result the commutator $[\phi(x),\phi^{\dagger}(y)]$ would have non zero contributions which must cancel outside the light cone to maintain causality. The term $\phi(x)\phi^{\dagger}(y)$ represents the propagation of a negatively charged particle from y to x and $\phi^{\dagger}(y)\phi(x)$ represents propagation of a positively charged particle from x to y. For both of these to exist

we would need these particles to have the same mass and so these cancel out. The causality in quantum field theory requires that for every particle there is a corresponding antiparticle which has the same mass but opposite charge. When the Klein Gordan field is real, a particle is it's own antiparticle.
If $\phi(x)$ is complex we can write it as $\phi(x) = (\phi_1(x) + \phi_2(x)\iota)/\sqrt{2}$ where $\phi_1(x)$ and $\phi_2(x)$ are both real valued functions. Since $\phi(x)$ is a solution, this would require both $\phi_1(x)$ and $\phi_2(x)$ must be solutions since $(\partial_\mu\partial^\mu + m^2)\phi_i(x)$ would be real hence we equate both to zero. Since $\phi^\dagger(x) = (\phi_1(x) - \phi_2(x)\iota)/\sqrt{2}$ this means we have

$$(\partial_\mu\partial^\mu + m^2)\phi(x) = 0 \tag{97}$$

$$(\partial_\mu\partial^\mu + m^2)\phi^\dagger(x) = 0 \tag{98}$$

Since $\phi_1(x)$ and $\phi_2(x)$ are two seperate solutions for the equation that are real valued we can add their contributions to the Lagrangian density as follows

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}(\partial_\mu\phi_1\partial^\mu\phi_1 + \partial_\mu\phi_2\partial^\mu\phi_2) - \frac{m^2}{2}(\phi_1^2 + \phi_2^2) \\ &= \partial_\mu\phi^\dagger\partial^\mu\phi - m^2\phi^\dagger\phi\end{aligned} \tag{99}$$

Similarly the Hamiltonian density would add up and we would have

$$\begin{aligned}\mathcal{H} &= \frac{1}{2}(\pi_1^2 + \pi_2^2) + \frac{1}{2}(\partial_\mu\phi_1\partial^\mu\phi_1 + \partial_\mu\phi_2\partial^\mu\phi_2) + \frac{m^2}{2}(\phi_1^2 + \phi_2^2) \\ &= \pi^\dagger\pi + \partial^\mu\phi^\dagger\partial_\mu\phi + m^2\phi^\dagger\phi\end{aligned} \tag{100}$$

Due to having two real valued fields as the solutions, we would have two sets of ladder operators for each momentum value. We will call them as $a_{\mathbf{p}}^{(1)}$ and $a_{\mathbf{p}}^{(2)}$. Using them we define $a_{\mathbf{p}} = (a_{\mathbf{p}}^{(1)} + a_{\mathbf{p}}^{(2)}\iota)/\sqrt{2}$ and $b_{\mathbf{p}} = (a_{\mathbf{p}}^{(1)} - a_{\mathbf{p}}^{(2)}\iota)/\sqrt{2}$ which lets us define

$$\phi(\mathbf{x}) = \int \frac{d^3}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\mathbf{p}}}}\left(a_{\mathbf{p}}e^{\iota\mathbf{p}\cdot\mathbf{x}} + b_{\mathbf{p}}^\dagger e^{-\iota\mathbf{p}\cdot\mathbf{x}}\right) \tag{101}$$

Some important points we can note here is that the two seperate solutions have the same mass hence representing the whole anti particle idea that was briefed upon in the previous section.

5.4 Detour: Jaynes-Cumming model

The Jaynes cumming model was introduced to explain the interaction of an atom with an electromagnetic field. The problem of an atom interacting with an electromagnetic had received semi-classical treatments which had however failed to explain all the characteristics of such a system. It received a proper quantum treatment where the field was quantized in the Jaynes-Cummings Hamiltonian [9]. This succesfully explained very important characteristics of the system like Rabi oscillations [12] and the collapse and revival of the system [6].
The field is treated as photons in a cavity where each particle is essentially treated as an excitation in this field. One can think of this as a Klein Gordan field but all the $\omega_{\mathbf{p}}$ is set to only one value effectively just making this into a normal SHO. The full Jaynes-Cummings Hamiltonian is given as follows: (ignoring the zero-point energy)

$$\begin{aligned}\mathbf{H}_{JC} &= \mathbf{H}_A + \mathbf{H}_F + \mathbf{H}_{A-F} \\ &= \hbar\omega_c\hat{a}_c^\dagger\hat{a}_c + \hbar\omega_{eg}|e\rangle\langle e| + \hbar g_c(\hat{\sigma}_+\hat{a_c} + \hat{\sigma}_-\hat{a}_c^\dagger)\end{aligned} \tag{102}$$

Here $\hat{a}_c$ is the ladder operation for the cavity field and $\sigma_\pm$ are the ladder operators of the two level atom ($|e\rangle, |g\rangle$). Shown below are some state evolution simulations of this system done using QuTip.

6 Feynman Diagrams

6.1 Path integral formulation

If one places a source in front of a wall with two slits, we will receive the very familiar interference pattern of the double slit problem. If we do happen to measure the screen however, we will see the pattern collapse to just two concentrated places as per the slit placement. This is pretty much the origin of the famous measurement problem but we will not concern ourselves with the elaborate history of measurement in quantum mechanics.
Now one can make the problem a wee bit more complicated considering n slots. One can use Kirchoff's diffraction and get an idea of the distribution formed on the other side using Fraunhoffer approximation. However things may feel messy if we have multiple screens each having any number of holes. Also one could consider every segment in

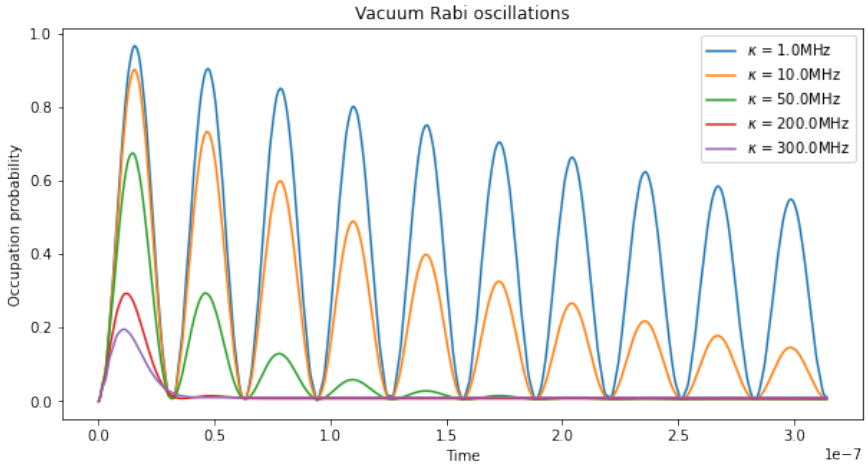

Figure 9: Shown above is a simulation of a two level atom which has been entangled with a cavity with 16 fock states which starts off in the first excited state. QuTip has been used for the simulation. The value of the coupling $g = 100$ MHz, the decay for the atom is $\gamma = 3$ MHz and the cavity decay is κ which is varying.

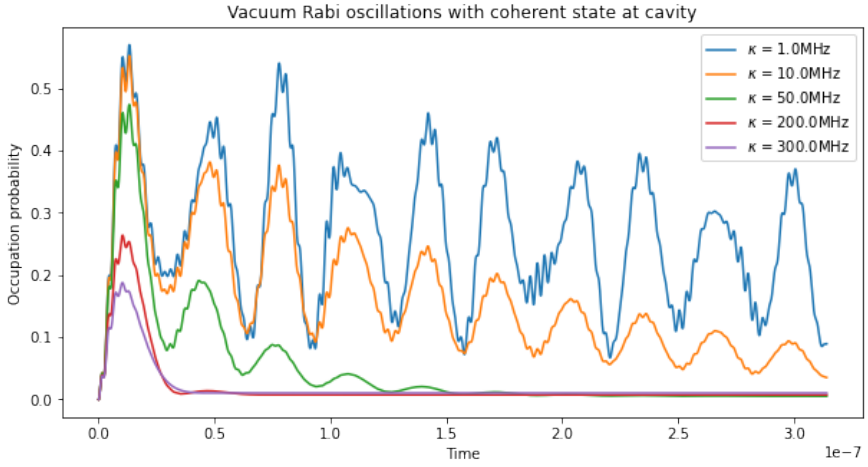

Figure 10: Shown above is a simulation of a two level atom which has been entangled with a cavity with 16 fock states which starts off in the coherent state with $\alpha = 1$. The value of the coupling $g = 100$ MHz, the decay for the atom is $\gamma = 3$ MHz and the cavity decay is κ which is varying.

space to be a screen which has holes all over it so is empty and what would one do then?
We now introduce the Dirac's formulation to understand this problem. We can generalize the problem as the probablity a particle reaches some point O in space from a source S in time T. We can label O as q_F and S as q_I and the propagation amplitude can be written as $\langle q_F | e^{-\iota H t} | q_I \rangle$. We can divide the evolution into N steps where each time-step is δt and each intermediate point can be multiplied over every possible path summed over. This would become into the following expression

$$\langle q_F | e^{-\iota H t} | q_I \rangle = \left(\prod_{j=1}^{N-1} \int dq_j \right) \langle q_F | e^{-\iota H \delta t} | q_{N-1} \rangle \langle q_{N-1} | e^{-\iota H \delta t} | q_{N-2} \rangle \dots \langle q_1 | e^{-\iota H \delta t} | q_I \rangle \tag{103}$$

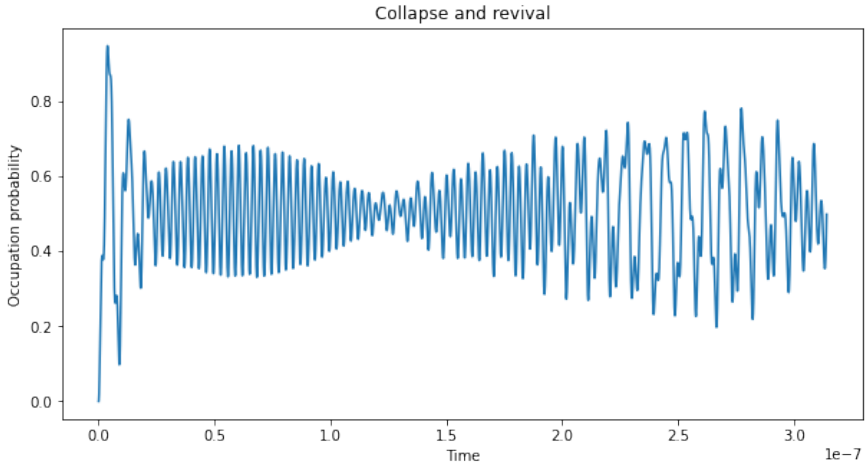

Figure 11: Demonstration of collapse and revival for a cavity with 32 fock states which starts off in the coherent state of $\alpha = 4$. As we can see the probability dies down and goes to 0.5 a little after 0.1μs and then again rises and oscillates.

This expression does not feel very inviting to be evaluated so let's start with the Hamiltonian of a free particle $H = p^2/2m$ and just look at one of the integrals

$$\langle q_{j+1}| e^{-\iota H\delta t(\hat{p}^2/2m)} |q_j\rangle = \int \frac{dp}{2\pi} \langle q_{j+1}| e^{-\iota H\delta t(\hat{p}^2/2m)} |p\rangle\langle p|q_j\rangle \tag{104}$$

$$= \int \frac{dp}{2\pi} e^{-\iota\delta t(p^2/2m)} e^{\iota p(q_{j+1}-q_j)} \tag{105}$$

$$= \left(\frac{-\iota m}{2\pi\delta t}\right)^{1/2} e^{\iota\delta t(m/2)[(q_{j+1}-q_j)/\delta t]^2} \tag{106}$$

We get this since it is a Gaussian integral. We can put this in the expression of equation 103 and we would get

$$\langle q_F| e^{-\iota Ht} |q_I\rangle = \left(\frac{-\iota m}{2\pi\delta t}\right)^{N/2} \left(\prod_{k=1}^{N-1} \int dq_k\right) e^{\iota\delta t(m/2)\sum_{j=0}^{N-1}[(q_{j+1}-q_j)/\delta t]^2} \tag{107}$$

In the limit of $N \to \infty$ we would replace $[(q_{j+1} - q_j)/\delta t]^2$ by $\dot{q}^2$ since the time interval is now infinitesimal and $\delta t \sum_{j=0}^{N-1}$ by $\int_0^T dt$ and we define the integral over paths as

$$\int Dq(t) = \lim_{N\to\infty} \left(\frac{-\iota m}{2\pi\delta t}\right)^{N/2} \left(\prod_{k=1}^{N-1} \int dq_k\right) \tag{108}$$

Thus we obtain the expression $\langle q_F| e^{-\iota Ht} |q_I\rangle = \int Dq(t) \exp\left(\iota \int_0^T dt[m\dot{q}^2/2 - V(q)]\right)$ from which we can recognize that the integral in exponent is just the Lagrangian and so it's integral is the action. So we write

$$\langle q_F| e^{-\iota Ht} |q_I\rangle = \int Dq(t) e^{\iota \int_0^T L(q,\dot{q})} \tag{109}$$

We now define $Z = \langle 0| e^{-\iota Ht} |0\rangle$ and as per the definition in the above equation we can easily generalize this to more number of particles by simply having $q_1, \ldots, q_N$. The Lagrangian would be used of this many particle system. Now let us take the example of a mattress of springs in which case the potential is somewhat defined as

$$V(q_1, q_2, \ldots, q_N) = \sum_{ab} \frac{k_{ab}}{2})(q_a - q_b)^2 + \ldots \tag{110}$$

Taking the continuum limit of this mattress spacing $l \to 0$ we can replace the label on a particle by a position vector $\vec{x}$ so we write $q(t, \vec{x})$ instead of $q_a(t)$. Here on out we would replace q by φ and call this as our field. We will now take

the liberty of taking $c = 1$ which we pretty much did in the previous section too.
The way we have defined the action, we would require it to be Lorentz invariant if we are to concern ourselves with special relativity. We must note that for the mattress we spoke of $(q_a - q_b)^2 \approx l^2(\partial\varphi/\partial x)^2$ which when we write as an action we have Lorentz invariant action to be

$$S = \int d^d x \left[\frac{1}{2}(\partial\varphi)^2 - \frac{1}{2}m^2\varphi^2 - \frac{g}{3!}\varphi^3 - \frac{\lambda}{4!}\varphi^4 + \dots\right] \tag{111}$$

Here d is the dimension which is often written as a $d = D + 1$ dimensional spacetime with time being the 1. A $0 + 1$ dimensional theory is just the regular quantum mechanics. We have already covered the Klein-Gordan and defined the propagators in it's context and this will be pivotal for the Feynman diagrams. Now suppose we have some kind of current in our space which creates excitation in our vaccum, we can account for this in the potential as some $J(x)\varphi(x)$ and so our final path integral is

$$Z = \int D\varphi e^{\iota \int d^4x[(\partial\varphi)^2/2 - V(\varphi) + J(x)\varphi(x)]} \tag{112}$$

6.2 The baby and the child

We start with what [17] refers to as a baby problem which is to evaluate the following integral

$$Z(J) = \int_{-\infty}^{\infty} dq e^{-\frac{1}{2}m^2q^2 - \frac{\lambda}{4!}q^4 + Jq} \tag{113}$$

When $\lambda = 0$ this just becomes into a Gaussian integral which can be easily calculated by completing square in the exponent. Now for starters we can try to expand the integral $Z(J)$ as a series in λ to get

$$Z(J) = \int_{-\infty}^{\infty} dq e^{-\frac{1}{2}m^2q^2 + Jq}\left[1 - \frac{\lambda}{4!}q^4 + \frac{1}{2}\frac{\lambda^2}{(4!)^2}q^8 + \dots\right] \tag{114}$$

One can use the trick of writing $\int_{-\infty}^{\infty} dq e^{-\frac{1}{2}m^2q^2 + Jq}q^4 n$ as some simply $\left(\frac{d}{dJ}\right)^{4n}\int_{-\infty}^{\infty} dq e^{-\frac{1}{2}m^2q^2 + Jq}$ and hence we get

$$\begin{aligned} Z(J) &= \left(1 - \frac{\lambda}{4!}\left(\frac{d}{dJ}\right)^4 + dfrac\lambda^2 2 \times (4!)^2\left(\frac{d}{dJ}\right)^8 + \dots\right)\int_{-\infty}^{\infty} dq e^{-\frac{1}{2}m^2q^2 + Jq} \\ &= e^{-\frac{\lambda}{4!}\left(\frac{d}{dJ}\right)^4}\int_{-\infty}^{\infty} dq e^{-\frac{1}{2}m^2q^2 + Jq} = \left(\frac{2\pi}{m^2}\right)^{1/2} e^{-\frac{\lambda}{4!}\left(\frac{d}{dJ}\right)^4} e^{J^2/2m^2} \end{aligned} \tag{115}$$

By expanding these two exponential series we can obtain any term in the double series expansion of $Z(J)$ in λ and J. We will be supressing the multiplicative constant by denoting it as $Z(0, 0)$ and define $\tilde{Z} = Z(J)/Z(0, 0)$. Say we want the term which has $\lambda^2 J^6$ we simply would have $\frac{1}{2}(\lambda/4!)^2(d/dJ)^8[1/(7!(2m^2)^7)]J^14$ and we can similarly construct many such examples. Now there will be certain pattern to each of these terms. We can associate diagrams to each term and make some rules on these basis and we arrive at the following rules

1. Diagrams are made of lines and vertices at which four lines meet
2. Each vertex is assigned a factor of $(-\lambda)$
3. Each line is assigned $1/m^2$
4. Assign a J for each external end of a line

We can note that twice the number of lines is equal to four times the number of vertices plus the number of ends which can clearly arrive from the way that each λ brings along a four time derivative of J.
We can also have chosen to start off with expanding Z in just terms of J and we would define

$$Z(J) = \sum_{s=0}^{\infty}\frac{J^s}{s!}\int_{-\infty}^{\infty} dq e^{-\frac{1}{2}m^2q^2 - (\lambda/4!)q^4}q^s = Z(0, 0)\sum_{s=0}^{\infty}\frac{J^s}{s!}G^{(s)} \tag{116}$$

The coefficients $G^{(s)}$ would be series in λ and are analogous to Green's functions in field theory.
We can also note that certain diagrams are connected and some are not. We can write the expression of $Z(J, \lambda)$ as

$$Z(J, \lambda) = Z(J = 0, \lambda)\sum_{N=0}^{\infty}\frac{1}{N!}[W(J, \lambda)]^N \tag{117}$$

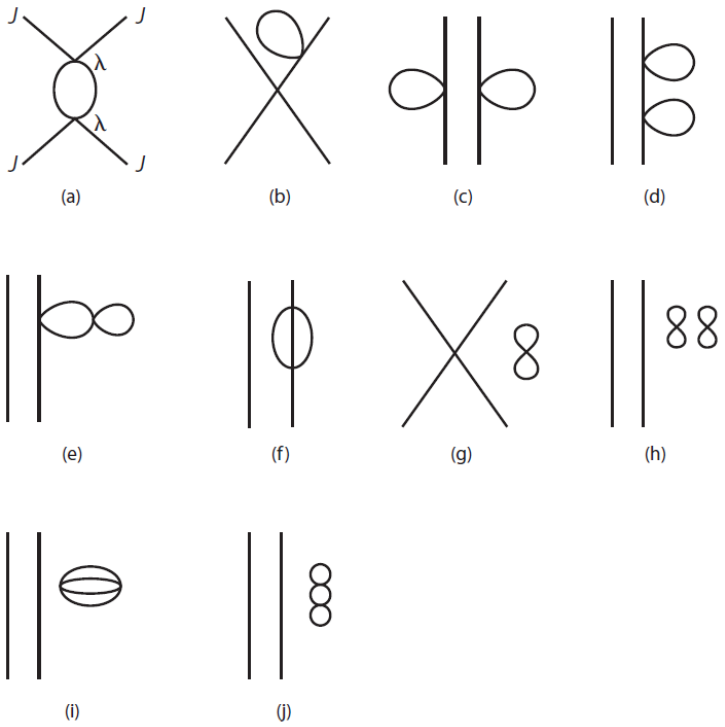

Figure 12: The diagrams for the term $\left(\frac{1}{m^2}\right)\lambda^2 J^4$ taken from [2]

By definition $Z(J=0,\lambda)$ represents diagrams which have no external source J. The statement is that W is a sum of connected diagrams while Z contains connected as well as disconnected diagrams. Say we happen to have n disconnected pieces in the diagram, those would be coming from $(1/n!)[W(J,\lambda)]^n$.

Now as the baby problem grows up to a child problem where q is now some multivariable term and not one dimensional. We can generalize and arrive at a similar expression of

$$\begin{aligned} Z(J) &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} dq_1\ldots dq_N e^{-\frac{1}{2}q\cdot A\cdot q-(\lambda/4!)q^4+J\cdot q} \\ &= \left[\frac{(2\pi)^N}{\det[A]}\right]^{1/2} e^{-(\lambda/4!)\sum_i(\partial/\partial J_i)^4} e^{\frac{1}{2}J\cdot A^{-1}\cdot J} \end{aligned} \tag{118}$$

A feature that this child problem has which the baby problem doesn't is the feature of propagation. We can also alternatively expand the integral in terms of power series for each component of J as follows

$$Z(j) = Z(0,0)\sum_{s=0}^{\infty}\sum_{i_1=1}^{N}\cdots\sum_{i_s=1}^{N} J_{i_1}\ldots J_{i_s}G^{(s)}_{i_1\ldots i_s} \tag{119}$$

We can think of the labeling i to be like labeling on a lattice. We would call $G^{(2)}_{ij}$ as the 2-point Green's function. We can see that $G^{(2)}_{ij}(\lambda=0) = (A^{-1})_{ij}$. The matrix term of $(A^{-1})_{ij}$ essentially represents the propagation from i to j. On evaluating the 4-point Green's function we get

$$G^{(4)}_{ijkl} = (A^{-1})_{ij}(A^{-1})_{kl} + (A^{-1})_{ik}(A^{-1})_{jl} + (A^{-1})_{il}(A^{-1})_{jk} - \lambda\sum_n (A^{-1})_{in}(A^{-1})_{jn}(A^{-1})_{kn}(A^{-1})_{ln} + \mathcal{O}(\lambda^2) \tag{120}$$

The first three terms represent an excitation propagation from i to j and another from k to l and permutations of this transfer. The first $\mathcal{O}(\lambda)$ term shows propagation to a certain n from all the i,j,k,l hence much like the diagram of 4 J terms coming to intersect at a $-\lambda$. One can however have n to n terms too which have not been accounted above which would come from Wick contraction but these will be dealt with later.

6.3 Perturbative field theory

Now we return to the actual integral for $Z(J)$ which is restated below

$$Z(J) = \int D\varphi e^{\iota \int d^4x[\frac{1}{2}[(\partial\varphi)^2 - m^2\varphi^2] - (\lambda/4!)\varphi^4 + J\varphi]} \tag{121}$$

A major difference here is that instead of using the discretized q_i we have a continuous variable function φ. Aside from that it is the same as the child problem we saw previously discussed. The expression finally reduces to the following

$$Z(J) = Z(0,0)e^{(-\iota/4!)\lambda \int d^4w[\delta/\iota\delta J(w)]^4} e^{-(\iota/2)\int\int d^4x d^4y J(x)D_F(x-y)J(y)} \tag{122}$$

The propagator $D_F(x-y)$ has it's expression written in equation 96. We have set the dimension to be $d = 4$ here but if it were set to $d = 0$, then $D_F(x-y)$ collapses to $-1/m^2$. $J(x)$ corresponds to sources and sinks and one can expand $Z(J)$ in terms of J, each power corresponds to a process with that many particles. For example a process where two particles collide to produce two more, this has four total J powers. We would obtain similar green functions in the J expansion as seen before.

$$Z(J) = Z(0,0)\sum_{s=0}^{\infty} \frac{i^s}{s!} \int dx_1 \dots dx_s J(x_1) \dots J(x_s) G^{(s)}(x_1, \dots, x_s) \tag{123}$$

$$G^{(s)}(x_1, \dots, x_s) = \int D\varphi e^{\iota \int d^4x[\frac{1}{2}[(\partial\varphi)^2 - m^2\varphi^2] - (\lambda/4!)\varphi^4]} \varphi(x_1) \dots \varphi(x_s) \tag{124}$$

A further thing to note is that $G(x_1, x_2)$ is only dependent on $x_1 - x_2$ not on their separate values and similarly $G(x_1, x_2, x_3, x_4)$ is only dependent on the differences. This is essentially their transnational invariance. $G(x_1, x_2)$ at $\lambda = 0$ just equals the translation from x_1 to x_2 much like the child problem and we can also see that $G(x_1, x_2, x_{3,4})$ described the scattering of particles.

The strategy for setting up sources and sinks helps in seeing the propagation of the particle. Say we have two sources and two sinks, we essentially are looking for the term with four J hence $G(x_1, x_2, x_3, x_4)$. Similar to how we did the child problem, the $O(\lambda)$ term here would be $(-\iota\lambda)\int d^4w D_F(x_1 - w)D_F(x_2 - w)D_F(x_3 - w)D_F(x_4 - w)$. This just says that the interaction point for these sources and sinks could be anywhere hence the integral over all w but by the behavior of D_F we know that those not allowed causality would simply become zero.

We must note that working in momentum space is much more convenient here as one can think of a meson with momentum k_1 colliding with another of k_2 to produce two new mesons of momenta k_3 and k_4. Each spacetime propogator gives us

$$D(x_a - w) = \int \frac{d^4k_a}{(2\pi)^4} \frac{e^{\pm k_a(x_a - w)}}{k_a^2 - m^2 + \iota\epsilon} \tag{125}$$

We note that the dummy variable integration can have either a plus or a minus sign in the exponential as it's inconsequential. This leads us to integrating over w for the $\mathcal{O}(\lambda)$ for the term

$$\int d^4w D_F(x_1 - w)D_F(x_2 - w)D_F(x_3 - w)D_F(x_4 - w) = \int d^4w e^{-\iota(k_1 + k_2 - k_3 - k_4)w} = (2\pi)^4\delta(4)(k_1 + k_2 - k_3 - k_4) \tag{126}$$

This interaction could occur anywhere in spacetime and that fact translates into momentum conservation $k_1 + k_2 = k_3 + k_4$. As stated the signs are chosen by choice of representation but the interpretation is the same regardless of that choice.

6.4 The rules

Now I will list the celebrated momentum space Feynman rules for scalar field theory.

1. Draw the Feynman diagram using the logic of the previous rules.
2. Label each line with some momentum k and associate it with the propagater $\iota/(k^2 - m^2 + \iota\epsilon)$.
3. Each interaction vertex has a coupling of $-\iota\lambda$ associated with it and the momentum entering it must equal that which exits it.
4. Momenta associated with internal lines are to be integrated over with the measure $d^4k/(2\pi)^4$.
5. Based on the symmetries present in the diagram, there would be a factor to be multiplied to it.
6. External lines must not have an associated propagator.
7. The delta function for overall momentum conservation is considered understood.

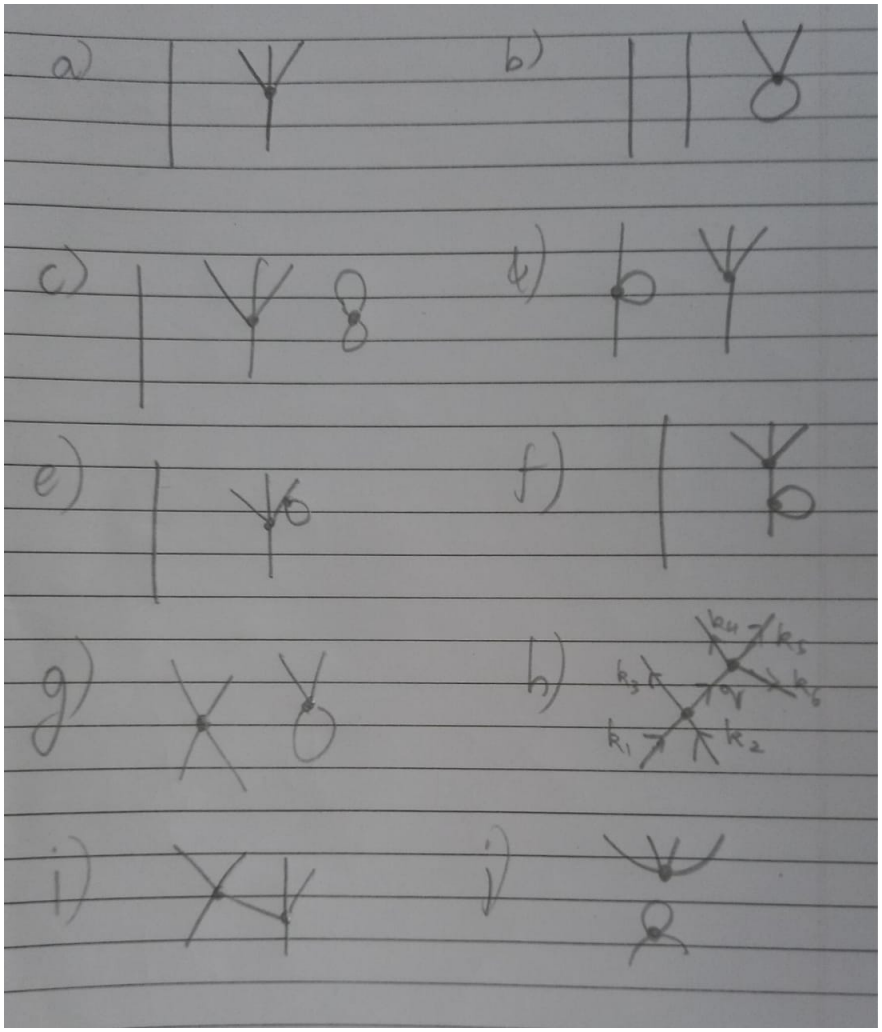

Figure 13: All the figures upto an $\mathcal{O}(\lambda^2)$ term for the formation of four mesons from two mesons. Not all figures are physically possible necessarily and the time moves upwards.

6.5 The birth of particles and loops

Let us have a look at the problem of two mesons colliding to create four mesons. Considering upto $\mathcal{O}(\lambda^2)$ terms we would be looking at $\lambda^2 J^6$ and λJ^6 terms for which we can draw the diagrams to be as shown in figure 13. Now let us take a look at figure 13(h) which arises from the λ^2 term. Here let us say that the internal line has some momentum q. Using rule 4 we can integrate over q to get

$$(-\iota\lambda)^2 \int \frac{d^4q}{(2\pi)^4} \frac{\iota}{q^2 - m^2 + \iota\epsilon} (2\pi)^4 \delta^{(4)}(k_1 + k_2 - k_3 - q)(2\pi)^4 \delta^{(4)}[q - (k_4 + k_5 + k_6)] \tag{127}$$

The integral over q us a cinch which gives

$$(-\iota\lambda)^2 \frac{\iota}{(k_4 + k_5 + k_6)^2 - m^2 + \iota\epsilon} (2\pi)^4 \delta^{(4)}[k_1 + k_2 - (k_3 + k_4 + k_5 + k_6)] \tag{128}$$

And this really just tells us that the overall momentum is conserved and as per rule 7 we wish to not drag this delta function around. The amplitude associated with this would be equal to the term preceding the delta function in the above expression. Now we can say that while it must have momentum equal to $k_4 + k_5 + k_6$, this need not be the same as m^2 and in this case it is penalized by having a significantly smaller amplitude hence the cost of not being real.
So far we have looked into some tree diagrams, however loop diagrams also do arise in the baby problem. As we can see in figures 14, 15 one can create such diagrams for λ^2 and λ^4 terms respectively. There however is a new funny

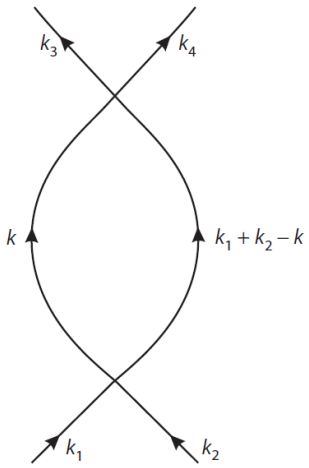

Figure 14: An example of a loop diagram with one loop, taken from [17]

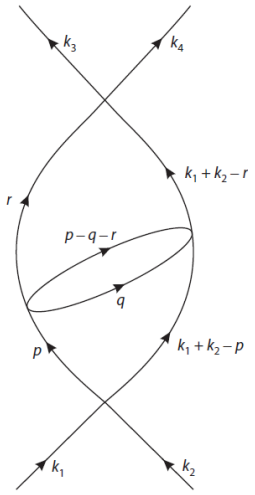

Figure 15: An example of a loop diagram with two loops, taken from [17]

behavior observed here where the integral happens to diverge. The form of the integral for 14 is

$$\frac{1}{2}(-\iota\lambda)^2\int\frac{d^4k}{(2\pi)^4}\frac{\iota}{k^2-m^\iota\epsilon}\frac{\iota}{(k_1+k_2-k)^2-m^2+\iota\epsilon} \tag{129}$$

Once again we can check that there will be cost for not being real much like the previous example. however for large k this integral essentially behave like $\int d^4k(1/k^4)$ which diverges, which makes it infinite and that cannot be a good thing. This is dealt with in renormalization which I shall not be describing in this report however. Similarly the triple integral for p, q, r of figure 15 will diverge as $\int d^{12}P(1/P^{12})$.

7 Dirac Field

7.1 Clifford Algebra

The motivation of the Dirac equation, as per a legend, arose when Dirac was one day staring into a fire. For some reason he wanted a wave equation that would be linear in ∂_μ. At first sight this may seem like this cannot be lorentz invariant since it would have some $c^\mu\partial_\mu$ acting on ψ to give something proportional to ψ where c^μ are constants. By the nature of c^μ, it would behave like a direction and so this would not be Lorentz invariant. Nevertheless, following modern notation the equation can be written as

$$(\iota\gamma^\mu\partial_\mu - m)\psi = 0 \tag{130}$$

On multiplying this equation with $(\iota\gamma^\mu\partial_\mu + m)$ one would obtain $-(\gamma^\mu\gamma^\nu\partial_\mu\partial_\nu + m^2)\psi = 0$. Since derivatives commute we can write $\gamma^\mu\gamma^\nu\partial_\mu\partial_\nu = \frac{1}{2}\{\gamma^\mu,\gamma^\nu\}\partial_\mu\partial_\nu$. From this we get that $\left(\frac{1}{2}\{\gamma^\mu,\gamma^\nu\}+m^2\right)\psi = 0$. Dirac noticed that if $\{\gamma^\nu.\gamma^\mu\} = 2\eta^{\mu\nu}$ this actually gives us the Klein-Gordan equation. In this situation one would note that for $\mu \neq \nu$ we have $\gamma^\nu\gamma^\mu = -\gamma^\nu\gamma^\mu$ hence these cannot be normal numbers.
Any numbers which satisfy $\{\gamma^\mu,\gamma^\nu\} = 2\eta^{\mu\nu}$ are said to form a Clifford algebra. One can check that the following 4×4 matrices satisfy the condition.

$$\gamma^0 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} = I\otimes\sigma_3 \tag{131}$$

$$\gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} = \sigma^i\otimes\iota\sigma_2 \tag{132}$$

We call these the gamma matrices. The Dirac equation in momentum space can write it as $(\gamma^\mu p_\mu - m)\psi(p) = 0$. More importantly, this equation is Lorentz invariant. In the rest frame this essentially becomes $(\gamma^0-1)\psi = 0$ and since $(\gamma^0-1)^2 = -2(\gamma^0-1)$ hence it somewhat behaves like a projection operator. This will when explicitly written

down, simply tell that 2 out of the 4 components of ψ would be zero. This makes sense as the electron has 2 physical degrees of freedom and not 4. A convenient notation which we will use from this point on is $\not{a} = \gamma^\mu a_\mu$ and so we write the Dirac equation as $(\iota\not{\partial} - m)\psi = 0$.
Now we must also define a form of a Lorentz transform for the state ψ similar to $x'^\nu = \Lambda^\nu_\mu x^\mu$. Given that ψ has four components it would transform as some $\psi'(x') = S(\Lambda)\psi(x)$ where $S(\Lambda)$ is a 4×4 matrix. To parametrize this we would need 16 linearly independent 4×4 matrices. We already have 5 of them, being the identity and the four gamma matrices. Since the square of γ^μ is equal to ± 1 and they anticommute with each other, the only matrices to consider would be of form $\gamma^\mu\gamma^\nu$, $\gamma^\mu\gamma^\nu\gamma^\lambda$ and $\gamma^\mu\gamma^\nu\gamma^\rho$ where all the indices are unequal. So we only have a single four gamma matrix product which we will write as $\gamma^5 = \iota\gamma^0\gamma^1\gamma^2\gamma^3$. This gives us

$$\gamma^5 = (I \otimes \sigma_3)(\sigma^1\sigma^2\sigma^2 \otimes \sigma_2) = I \otimes \sigma_1 \tag{133}$$

An important thing to note is that $\{\gamma^5, \gamma^\mu\} = 0$. We also get four new products which are $\gamma^\mu\gamma^5$ which are really just the same as taking three distinct gamma matrices in conjunction. Finally we can see that $\gamma^\mu\gamma^\nu = \eta^{\mu\nu} - \iota\sigma^{\mu\nu}$ where $\sigma^\mu\nu = \frac{\iota}{2}[\gamma^\mu, \gamma^\nu]$. Here we have 6 such matrices. We can obtain the expressions as follows of $\sigma^{\mu\nu}$

$$\sigma^{0i} = \iota \begin{pmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{pmatrix} \tag{134}$$

$$\sigma^{ij} = \epsilon^{ijk} \begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix} \tag{135}$$

This finally gives our nice set of 16 matrices which is $\{1, \gamma^\mu, \sigma^{\mu\nu}, \gamma^\mu\gamma^5, \gamma^5\}$. Now we can begin to define the Lorentz transformation using this set.
As we explored in quantum mechanics, general rotations are represented as $e^{\iota\vec{\theta}\vec{J}}$ where $\vec{J}$ are the generators of rotation. We can in similar notation describe the Lorentz transformation as $\Lambda = e^{-\frac{\iota}{2}\omega_{\mu\nu}J^{\mu\nu}}$ where $\omega_{\mu\nu}$ is assymetric and so the 6 components describe 3 boosts and 3 rotations and J^{0i} generating boosts and J^{ij} generating rotations. More specifically we note that J^{ij} is represented by $\sigma^{ij}/2$. This fits the description for rotation of a spin-1/2 particle taking the first two and last two components of ψ individually.
We finally arrive at $S(\Lambda) = e^{-(\iota/4)\omega_{\mu\nu}\sigma^{\mu\nu}}$ and $\psi'(x') = S(\Lambda)\psi(x)$ and this satisfies the equation $(\iota\gamma^\mu\partial'_\mu - m)\psi'(x') = 0$. Similar to how quantum mechanical rotations are built we can note that for infinitesimal ω we have $S\gamma^\lambda S^{-1} = \gamma^\lambda - (\iota/4)\omega_{\mu\nu}[\sigma^{\mu,\nu}, \gamma^\lambda] = \gamma^\lambda + \gamma^\mu\omega^\lambda_\mu$. Building up to a finite Lorentz transform we get that $S\gamma^\mu S^{-1} = \Lambda^\lambda_\mu\gamma^\mu$.
We know that the Clifford algebra gives $(\gamma^0)^2 = 1$ with γ^0 hermitian and $(\gamma^i)^2 = -1$ with γ^i anti-hermitian and so one can conveniently express $(\gamma^\mu)^\dagger = \gamma^0\gamma^\mu\gamma^0$. This leads us to realising that the bilinear $\psi^\dagger\gamma^\mu\psi$ is not hermitian. However $\tilde{\psi}\gamma^\mu\psi$ is where $\tilde{\psi} = \psi^\dagger\gamma^0$. We can also from this trivially note that as per the transformation rule we would have $\tilde{\psi}'(x')\psi'(x') = \tilde{\psi}(x)\psi(x)$

7.2 $\mathcal{CPT}$ symmetry and more

A very common discrete symmetry in physics is that of parity which is reflection across a mirror. We can reach an arbitrary reflection by adding a rotation to the transform $x'^\mu = (x^0, -\vec{x})$. The parity here would be obtained by the transform of γ^0 which trivially does not affect the actual solution of the equation so we can just write $\gamma'(x') = \kappa\gamma^0\psi(x)$ where κ is some arbitrary phase.
This now leads us to discussing the actual Lagrangian of the system which actually happens to be

$$\mathcal{L} = \tilde{\psi}(\iota\not{\partial} - m)\psi \tag{136}$$

We can check that the Euler-Lagrange equations on this give the equation $\partial_\mu(\iota\tilde{\psi}\gamma^\mu) + m\tilde{\psi} = 0$ which upon conjugation and multiplication from the right with γ^0 gives the good old Dirac equation.
Given some set of gamma matrices we can solve the Dirac equation as $(\not{p} - m)\psi(p) = 0$. Now one could change the basis and use $\gamma'^\mu = W^{-1}\gamma^\mu W$ and this should not change the physics. The choice of a convenient basis would vary in cases.
For slowly moving electrons we can pick the normal basis already defined in equations 131 and 132. The 2-component description of $\psi = [\chi \quad \phi]^T$ where each of χ and ϕ are two components would give a χ much smaller than ϕ.
In contrast when the momentum is much larger one can approximate the equation to just be $\not{p}\psi(p) = 0$. On multiplying from the left by γ^5 we can claim that $\gamma^5\psi$ is also a solution as this operator anti-commutes with all the other gamma matrices. Since $\gamma^5 = 1$ we can form two projection operators $P_L = \frac{1}{2}(1 - \gamma^5)$ and $P_R = \frac{1}{2}(1 + \gamma^5)$ which satisfies idempotency and $P_L P_R = 0$. Also introducing $\psi_L = \frac{1}{2}(1 - \gamma^5)\psi$ and $\psi_R = \frac{1}{2}(1 + \gamma^5)\psi$ will make our lives easier

and additionally we can note that $\gamma^5\psi_L = -\psi_L$ and $\gamma^5\psi_R = \psi_R$ hence are eigenstates for γ^5. The fact that ψ_R and ψ_L are eigenstates for γ^5 means that working in a basis where γ^5 is diagonal would help us out if we write terms in ψ_R, ψ_L.
Using this we define two basis, one being the Weyl basis where γ^5 is diagonal and the other being the Dirac basis where γ^0 is diagonal. This now leads us to the chirality of the field. Using these terms defined we can rewrite the Lagrangian as follows

$$\mathcal{L} = \tilde{\psi}_L \iota \not{\partial} \psi_L + \tilde{\psi}_R \iota \not{\partial} \psi_R - m(\tilde{\psi}_L \psi_R + \tilde{\psi}_R \psi_L) \tag{137}$$

The interesting part here is that if one were to transform $\psi \to e^{\iota\theta}\psi$ the Lagrangian $\mathcal{L}$ is invariant to it. Since $\psi = \psi_R + \psi_L$ the phase would apply on both the right and left handed fields (the names given to ψ_R and ψ_L respectively). As per Noether's theorem we know this invariance implies existence of some conserved current. Here it turns out to be $J^\mu = \tilde{\psi}\gamma^\mu\psi$ which feels fairly trivial.
If m were to be zero, the transformation $\psi \to e^{\iota\theta\gamma^5}\psi$ would also be invariant on the Lagrangian and the current $J^{5\mu} = \tilde{\psi}\gamma^\mu\gamma^5\psi$. This is known as the chiral symmetry as the left and right handed fields transform as $\psi_L \to e^{-\iota\theta}\psi_L$ and $\psi_R \to e^{\iota\theta}\psi_R$ respectively.
On a detour we can introduce a field φ which is interacting to the Dirac field. This would amount to adding some term $g\varphi\tilde{\psi}\psi$ to the Lagrangian where g is just a coupling constant. Similarly for a vector fiedl we would add the term $eA_\mu\tilde{\psi}\gamma^\mu\psi$ and also in this case we can introduce a covariant derivative $D_\mu = \partial_\mu - \iota e A_\mu$. This lets us write the Lagrangian $\mathcal{L} = \tilde{\psi}(\iota\not{\partial} - m)\psi + eA_\mu\tilde{\psi}\gamma^\mu\psi = \tilde{\psi}(\iota\gamma^\mu D_\mu - m)\psi$. Hence the Lagrangian for a Dirac field interacting with a vector field of mass μ is

$$\mathcal{L} = \tilde{\psi}(\iota\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{\mu^2}{2}A_\mu A^\mu \tag{138}$$

When the mass μ vanishes, this becomes the Lagrangian for quantum electrodynamics. Varying with respect to $\tilde{\psi}$ we get the Dirac equation in presence of an electromagnetic field as

$$[\iota\gamma^\mu(\partial_\mu - \iota e A_\mu) - m]\psi = 0 \tag{139}$$

Now we move on to charge conjugation. We note that taking the complex conjugate of the above equation leads us to $[-\iota\gamma^{\mu *}(\partial_\mu + \iota e A_\mu) - m]\psi^* = 0$ which can be thought of as flipping the charge. Also $-\gamma^{\mu *}$ happens to satisfy the Clifford algebra so really it just is a different basis hence we can say that there is some matrix $C\gamma^0$ such that $-\gamma^{\mu *} = (C\gamma^0)^{-1}\gamma^\mu(C\gamma^0)$ which on plugging in we can write

$$[\iota\gamma^\mu(\partial_\mu + \iota e A_\mu) - m]\psi_c = 0 \tag{140}$$

Where $\psi_c = C\gamma^0\psi^*$. This gives us that ψ_c us essentially the field of a particle which has the same mass as the electron but the opposite charge, namely the positron. We can more specifically look into how the matrix C must be defined and that it requires $C\gamma^0\gamma^{\mu *}\gamma^0 C^{-1} = -\gamma^\mu$ and we can reduce this to $(\gamma^\mu)^T = -C^{-1}\gamma^\mu C$ if γ^0 is a real matrix.
In both the Weyl and Dirac basis, the only imaginary matrix is γ^2 hence $C\gamma^0$ essentially anticommutes with all gamma matrices except γ^2 with which it commutes as per the definition. So we essentially end up with $C\gamma^0 = \gamma^2$ barring some phase factor. This means that $\psi_c = \gamma^2\psi^*$. This also shows that the handedness of this conjugate would be the opposite of that of ψ based on how the γ^5 operator acts on them.
So well we have the $\mathcal{C}$ (charge conjugation) and $\mathcal{P}$ (parity) of our nice looking $\mathcal{CPT}$ somewhat developed so let's move on to $\mathcal{T}$ (time reversal). In regular quantum mechanics, the time reversal operator is one which is anti-hermitian which simply can be proven by writing out the evolution. This is essentially the transformation $t' = -t$ and we wish to find $\iota(\partial/\partial t)\psi'(t') = H\psi'(t')$. We would write $\psi'(t') = T\psi(t)$ which makes $\iota[\partial/\partial(-t)]T\psi(t) = HT\psi(t)$. Since H has no time dependence in this case, on multiplying T^{-1} we must note that H and T^{-1} should commute. This leads us to a somewhat odd conclusion that $T^{-1}(-\iota)T = \iota$. This lets us define $T = UK$ where U is unitary and K is an operator which complex conjugates everything which essentially makes T anti-unitary.
One could act T on a plane wave where U would simply be a phase factor. In this case $\psi'(t)$ would have positive energy and just propogates in the other direction. Now on a spin-1/2 system we would want it to be able to flip spins and so one could for example just use $U = \eta\sigma_2$ where η is some phase. Importantly $T^2 = -1$ which is the Kramer's degeneracy's origin. This essentially states no matter how complicated, for any odd number of electrons in an electric field, each energy level will be two-level degenerate.
The proof takes the system to be time reversal invariant and so $T\psi = e^{\iota\alpha}\psi$ as they have the same energy. However $T^2\psi = T(T\psi) = Te^{\iota\alpha}\psi = e^{-\iota\alpha}T\psi = \psi \neq -\psi$ as one must expect if they were the same state, implying that they are a different state.
Getting to the Dirac equation, we can write it's Hamiltonian as $H = -\iota\gamma^0\gamma^i\partial_i + \gamma^0 m$ which paired with $T^{-1}HT = H$ gives us the unitary to be $U^{-1}\gamma^0 U = \gamma^{0*}$ and $U^{-1}\gamma^i U = -\gamma^{i*}$. Restricting to the Weyl and Dirac bases, the clearcut guess we get is that $U = \eta\iota\sigma^2 \otimes I$. Now finally we reach

Theorem 7.1 ($\mathcal{CPT}$ theorem)**.** The $\mathcal{CPT}$ theorem states that under the combined action of charge conjugation, parity and time reversal, any local Lorentz invariant field theory must be invariant to this transformation.

While proving this for any general field theory is well beyond the scope of this SoS for me, we can see that with the developed theory, it can be seen to hold for the Dirac field.

7.3 The Dirac Sea

Now while the Dirac equation looks pretty nice and cool, there are a few issues with it for which the Dirac hole theory [5, 7] was somewhat of a patchwork. Now before we dive in this interesting theory, why do we even need it? Well it turns out that as the Dirac equation has been described, it has negative energy solutions which is not all that much of a nice look. What is more problematic is that one cannot simply ignore them since the positive energy states do not form a complete basis by themselves.
Now this immediately raises a new issue, why don't the electrons simply go to the lowest state by emitting some photon of energy $2mc^2$ as each state is associated with it's negative energy version too. However a more fundamental question here would be why must we even assume that the vacuum state is infact reflective of what we know to be vacuum, i.e., emptiness? This is where the Dirac sea comes into picture, Taking into account Pauli's exclusion principle to work in advantage, one can assume that the negative energy continuum is filled completely which now represents the vacuum and all the positive energy states are empty. This is in essence what the Dirac sea is, a vacuum state filled with an infinite (sea) of electrons
A vacuum must be unobservable by having zero mass and no charge but this charge has an infinitely large negative energy and charge so that is an issue. We hence would renormalize this to define the Dirac sea to have no mass and no charge and the net charge would be defined with respect to this Dirac sea. This leads us to introduce the new concept of a hole which is the absence of an electron. As per this new theory, a hole has the charge of $+e$. Similarly the hole has the opposite value for momentum.
Now given our definition of vacuum, we can note that it can be modified by influence of external fields. Quite similar to holes and electrons moving around in semiconductors in the theories for semiconductor physics we can see that there would be radiation based on electrons getting devoured by holes. Many were to point out that this formalism wasn't very satisfactory as this still leaves the Klein-Gordan field's negative energy a mystery as it works only with Pauli's exclusion principle. The more favorable interpretation is the Feynman-Stückelberg interpretation which treats the negative energy to be associated with a positron going backwards in time as we can see from the time reversal theory. However one can argue that indeed a sea of electrons does sound pretty intriguing.

8 Acknowledgements

I would like to thank my mentor Som Phene who guided me through this very vast field as based on my very fluctuating interests and also shared a lot of interesting concepts with me in discussion. While titled as Quantum Field Theory, I have managed to explore the math behind classical field theory and also the normal Klein-Gordan and Dirac fields of QFT and this has definitely made this a fun experience for me.

References

[1] Koro (127). *Poincare recurrence theorem*. [Online; accessed 08-June-2021]. 2013. URL: https://planetmath.org/poincarerecurrencetheorem.

[2] Ruggero Maria Santilli (auth.) *Foundations of Theoretical Mechanics I: The Inverse Problem in Newtonian Mechanics*. 2nd ed. Texts and Monographs in Physics. Springer-Verlag Berlin Heidelberg, 1978. ISBN: 978-3-662-23683-3,978-3-662-25771-5.

[3] Mark Alford. *Legendre Transforms*. 2019. URL: https://web.physics.wustl.edu/alford/physics/Legendre_introduction.pdf.

[4] H. B. G. Casimir and D. Polder. "The Influence of Retardation on the London-van der Waals Forces". In: *Phys. Rev.* 73 (4 Feb. 1948), pp. 360–372. DOI: 10.1103/PhysRev.73.360. URL: https://link.aps.org/doi/10.1103/PhysRev.73.360.

[5] Paul Adrien Maurice Dirac and Ralph Howard Fowler. "A theory of electrons and protons". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 126.801 (1930), pp. 360–365. DOI: 10.1098/rspa.1930.0013. eprint: https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1930.0013. URL: https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1930.0013.

[6] J. H. Eberly, N. B. Narozhny, and J. J. Sanchez-Mondragon. "Periodic Spontaneous Collapse and Revival in a Simple Quantum Model". In: *Phys. Rev. Lett.* 44 (20 May 1980), pp. 1323–1326. DOI: 10.1103/PhysRevLett.44.1323. URL: https://link.aps.org/doi/10.1103/PhysRevLett.44.1323.

[7] Douglas M. Gingrich. *Hole Theory*. [Online; accessed 10-July-2021]. 2004. URL: https://sites.ualberta.ca/~gingrich/courses/phys512/node63.html.

[8] Ryzhik. Gradshteyn. *Table of integrals, series, and products*. 5th ed. AP, 1996.

[9] E.T. Jaynes and F.W. Cummings. "Comparison of quantum and semiclassical radiation theories with application to the beam maser". In: *Proceedings of the IEEE* 51.1 (1963), pp. 89–109. DOI: 10.1109/PROC.1963.1664.

[10] Brian Kend. *Manifolds: A Gentle Introduction*. 2018. URL: https://bjlkeng.github.io/posts/manifolds/.

[11] Encyclopedia of Mathematics. *First Integral*. URL: http://encyclopediaofmath.org/index.php?title=First_integral&oldid=36265.

[12] I. I. Rabi. "Space Quantization in a Gyrating Magnetic Field". In: *Phys. Rev.* 51 (8 Apr. 1937), pp. 652–654. DOI: 10.1103/PhysRev.51.652. URL: https://link.aps.org/doi/10.1103/PhysRev.51.652.

[13] Harish Seshadari. *Short Talk-What is a Manifold-I*. Youtube. 2017. URL: https://www.youtube.com/watch?v=cLRtp5yytAg.

[14] K. Vogtmann V. I. Arnold A. Weinstein. *Mathematical Methods Of Classical Mechanics*. 2nd. Graduate Texts in Mathematics. Springer, 1989. ISBN: 9780387968902,0387968903,3540968903,9783540968900.

[15] Wikipedia contributors. *Liouville's formula*. [Online; accessed 08-June-2021]. 2021. URL: https://en.wikipedia.org/wiki/Liouville%5C%27s_formula.

[16] Wikipedia contributors. *Manifold*. [Online; accessed 09-June-2021]. 2021. URL: https://en.wikipedia.org/wiki/Manifold.

[17] A. Zee. *Quantum Field Theory in a Nutshell*. illustrated edition. Princeton University Press, 2003. ISBN: 9780691010199,0-69-1-01019-6.