

Quantum field theory II

Lectures by: Professor Petr Hořava

Notes by: Jackson Van Dyke; All errors introduced are my own.

Contents

1. Introduction	5
Chapter 1. Bosonization	6
1. History of bosonization	6
2. Bosonic field theory	6
3. Fermionic field theory	11
4. Equivalence	12
5. Massless Schwinger model	13
6. Thirring model - sin-Gordon theory	15
7. Theory classification	17
Chapter 2. Non-abelian Yang-Mills	19
1. Lie groups	19
2. Generalizing scalar fields	21
3. Quantization	24
4. Gauge fixing in YM	29
5. Feynman rules	33
6. BRST symmetry	38
Chapter 3. Generic gauge theories	40
1. Instantons	40
2. BRST quantization of gauge theories	45
3. Topological quantum field theories	48
4. Topological Yang-Mills gauge theories	49
5. Super-symmetry	52
6. Chern-Simons theory	58
7. BRST-BV Formalism	66
Chapter 4. Large N (AdS / CFT correspondence)	74
1. Historical context	74
2. 1 + 1 dimensions	75
3. Riemannian surfaces	77
4. Generic treatment	78
5. AdS space	80
6. Establishing the dictionary	82
7. Strength of the equivalence	85
8. Anomalies	85
Chapter 5. Effective field theory of gravity	92
1. Preliminaries	92

CONTENTS

	4
2. UV-IR connection?	92
3. Review of EFT	92
4. Gravity as an EFT	93
5. Effective field theory of inflation	95
Chapter 6. Applications of QFT to finance	99
1. Fluctuations of financial assets	99
2. Path integrals	103
Bibliography	105

1. Introduction

QFT is a calculus for describing systems with many interacting and fluctuating degrees of freedom. This covers high energy particle physics, where we have fluctuations due to \hbar , as well as condensed matter (where \hbar also plays a role), and classical systems in condensed matter where the temperature plays this role. In particular, QFT is most useful for describing local systems in the sense that the constituents interact with only their local counterparts. More succinctly, QFT is a machinery which takes a chosen set of symmetries and produces predictive results. Our goals will fall in the two categories:

- (1) Techniques and calculations of QFT
- (2) Big picture, looking towards the horizons of the field

In particular, this semester is meant to approach open questions, rather than solving problems solved by many people in the past. We will first focus on Gauge theories, and then move to effective field theories (EFT). In terms of methods, we will concern ourselves over methods of renormalization and the renormalization group (RG). The other main camp of methods consists of the so-called *non-perturbative* techniques. This is not entirely equivalent, but it is usually correlated with the role of topology and geometry in physics. There are of course overlaps between these four topics, but this is roughly what will be covered:

- (1) Gauge theories: QED, Ward identities, non-Abelian Yang-Mills (YM) theories (mostly in four relativistic dimensions), quantization of non-Abelian YM, Faddeev-Popov ghosts, BRST (as a generic technique for quantizing arbitrary gauge theories), BRST-BV (involving good aspects of cohomology theories), supersymmetry, topological QFT, topological YM gauge theory (as an example of BRST quantization), Chern-Simons gauge theory (mostly in $2 + 1$ dimensions), (fractional) quantum Hall effect(s), anomalies, spontaneous symmetry breaking, Higgs mechanism.
- (2) RG: Confinement of QED in $1 + 1$ dimensions, bosonization in 2D (relates free fields describing bosons, to free fields describing fermions), unitarity of S -matrix (optical theorem), Wilsonian renormalization, critical exponents, nonrelativistic novelties (apply the idea of RG to non-rel. field theories)(gravity)
- (3) Nonperturbative/topology and geometry: cohomology, dualities (between weak and strong coupling of two different theories), anomalies (alg. topology, diff. geometry) (local and global), vortices, monopoles, solitons, instantons, K-theory (fancy cohomology), topological insulators, stable Fermi surfaces, large N expansion (taking large limit of DOF, provides duality for ADS-CFT correspondence)^{0.1}
- (4) EFT: Effective actions, low energy QCD, chiral Lagrangian, effective field theory of gravity, effective field theory of cosmological inflation (Keldysh-Schwinger).

^{0.1} Whenever you have a chance to find a new duality, do so. For example if no one ever came up with local Gauge symmetries, we can write one down with no gauge symmetries, then you can consider a large number of degrees of freedom, and discover the local gauge symmetries this way. So one should always ask what the dual of some particular symmetry is.

CHAPTER 1

Bosonization

This is a good opportunity to review some concepts from the first half of this course. We can review these in the context of free field theories. Bosonization itself, is a beautiful equivalence between a class of bosonic field QFTs, and a class of fermionic QFTs. If you are studying a theory, this theory is not determined by the Lagrangian, but rather the available observables (operators on your Hilbert space). As such, if someone else is studying another theory, then if there is a mathematical isomorphism between such observables, these theories are somehow the same. This notion is called *duality*.

We will see this in $1 + 1$ relativistic dimensions and add in self interactions later. Here we are truly entering a non-perturbative regime, because we are not depending only on Feynman diagrams and weak coupling expansions. In particular, using techniques from before, we can start with no coupling, turn on small self-coupling, and then using this strong-coupling duality, we can switch to the dual description, and use the same weak coupling tools in this strongly coupled regime. When we turn on this strongly coupling duality for a free field theory, we get the so called sine-Gordon/Thirring model.^{1.1}

1. History of bosonization

The first real introduction of bosonization was from D.C. Mattis and E. Lieb in 1965 [19]. Another landmark paper in the field came from A. Luther and I. Peschel published in 1975 [16]. This was around the same time that S. Coleman and S. Mandelstam introduced the sine-Gordon model in [18]. also in 1975. Since then, this has developed into a very powerful tool. The most famous reference for the applications in condensed matter is Haldane's "Luttinger liquids" [11]. There are also many applications of bosonization to string theory.

2. Bosonic field theory

Take a scalar field $\varphi \in \mathbb{R}$. Start with an action

$$(1.1) \quad S = \int d^D x \left\{ \frac{1}{2} (\partial\varphi)^2 + \cancel{V(\varphi)} \right\}$$

where we are Wick rotating to Euclidean space. In other words, we are integrating $\mathcal{Z} = \int \mathcal{D}\varphi(x) e^{-S}$ rather than $\mathcal{Z} = \int \mathcal{D}\varphi(x) e^{iS}$. It is tempting to think this simple theory is boring, but exactly the opposite is true. We already know that the naive classical "engineering" dimension of the field is

$$(1.2) \quad [\varphi]_\ell = \frac{D-2}{2}$$

^{1.1} This theory is treated more thoroughly in [7].

Note this may require quantum corrections from loops and the like. This means $D = 2$ is special. To see this, let's calculate the two point function:

$$(1.3) \quad \langle \varphi(x) \varphi(x') \rangle = \frac{\Gamma\left(\frac{D-2}{2}\right)}{4\pi^{D/2}} \left(\frac{1}{x^2}\right)^{\frac{D-2}{2}}$$

so for $D \rightarrow 2$, this result becomes questionable. This motivates the notion that we need a UV regulator in this theory. We also see that we need to step back and introduce an IR regulator as well. These regulators lead to interesting results.

In momentum space our propagator can be written:

$$(1.4) \quad \int \frac{d^D k}{(2\pi)^D} \frac{e^{ikx}}{k^2} = \int_0^\infty ds \int \frac{d^D k}{(2\pi)^D} \exp(-sk^2 + ikx)$$

$$(1.5) \quad = \int_0^\infty ds \int \left(\sqrt{\frac{\pi}{s}}\right)^D \frac{e^{-\frac{x^2}{4s}}}{(2\pi)^D}$$

$$(1.6) \quad = \frac{1}{2^D \pi^{D/2}} \int_0^\infty ds s^{\frac{D}{2}-1} e^{-\frac{sx^2}{4}}$$

So we need to do something more sophisticated due to the issues at $D = 2$. In particular, we need to add an infrared regulating mass $m_0 \neq 0$. So the propagator in the case of $D = 2$, $G(m, m_0^2)$, depends on the square of this new mass.^{1,2} Now let's back up, and not be so exclusive with this theory. To generalize as much as possible, we go back and introduce this new mass to our action. This becomes:

$$(1.7) \quad S(\varphi) = \int d^2x \left\{ \frac{1}{2} (\partial\varphi)^2 + \frac{1}{2} m_0^2 \varphi^2 \right\}.$$

So now $\psi \in S^1$ and we have a new geometric parameter, R , the radius of S^1 . We will write $g \equiv R^{-1}$ to remind us this is a coupling constant. The change of coordinates here is somewhat up to us. We are on S^1 , so we have two obvious choices:

$$(1.8) \quad \varphi = \varphi + 2\pi \quad \psi = \psi + 2\pi R$$

We choose the first one. So where does the radius show up? As a coupling constant in front of the whole thing:

$$(1.9) \quad S(\varphi) = \frac{1}{g^2} \int d^2x \left\{ \frac{1}{2} (\partial\varphi)^2 + \frac{1}{2} m_0^2 \varphi^2 \right\}$$

we cannot absorb this into the action, so indeed g is a physical coupling constant^{1,3} An example of where this periodicity shows up is in composite operators. So for $V_\alpha := e^{i\alpha\varphi}$, for which values of α is this operator physical? Of course this is when α takes integer values. So our propagator is now

$$(1.10) \quad G(x, m_0^2) = g^2 \int \frac{d^2k}{(2\pi)^2} \frac{e^{ikx}}{k^2 + m_0^2}$$

$$(1.11) \quad \approx g^2 \left(-\frac{1}{2\pi} \log m_0 |x| + \text{const.} + \mathcal{O}(m_0 |x|) \right)$$

^{1,2} Note when you list coupling constants as arguments of the propagator, it is typical to list them in the order in which they appear in the action. This can play a roll in considerations of naturalness etc.

^{1,3} this is bad terminology, since it doesn't reflect coupling directly, but it is a parameter on which the system depends

so the obstacle as $m_0 \rightarrow 0$ is logarithmic. How does one deal with this? Before we answer this we consider some properties of this propagator.

The first conclusion from this propagator, is something people often find shocking. This is however a physical result with measurable consequences. The massless scalar field φ , in $1 + 1$ dimensions doesn't exist as a propagator. Any such interpretation doesn't make physical sense, because the endpoint functions depend on m_0 . An immediate important consequence, is that there is no spontaneous symmetry breaking ever in global internal continuous symmetries in $1 + 1$ relativistic dimensions. We will see this again in a more formal way. This statement is actually known as the Coleman^{1,4} theorem. How do we prove this? We just rely on Goldstone's theorem:

THEOREM 1.1. *In a relativistic QFT, if a continuous global internal symmetry is broken, this automatically implies a massless excitation φ in the system. If this does not exist in some particular dimension, then it cannot ever exist. So there couldn't have ever been a breaking of such a symmetry.*

So this theorem immediately implies that φ doesn't make sense as a physical object. Note here continuous means symmetries generated by infinitesimal generators. So not discrete symmetries. For example, a reflection of the circle is discrete, and Goldstone has nothing to say about this. Global means not gauge-symmetry. Internal means as opposed to spacetime symmetries. We might wonder if the same story applies to spacetime symmetries. Actually yes, but the theorem is actually slightly stronger, because there is a one to one correspondence between the number of this, and the number of broken symmetries. So internal is one implication, spacetime is biconditional.

So if φ doesn't exist, how can we claim this is a good theory? Instead, there is indeed a story for such a theory. This is ok because we don't care so much about the Lagrangian, we just want a list of observables (operators in the Hilbert space) with corresponding correlation functions. We can list some of them as composite operators. In particular, $\partial_\mu \varphi$ is a meaningful operator. It is actually very important. A shift symmetry by a constant (rotation) is a good classical symmetry, and cannot be broken, since we have no such spontaneous symmetry breaking. We will see quite explicitly, that this cannot be broken by calculating correlation functions directly. This is important because the conserved current of such a $U(1)$ symmetry is:

$$(1.12) \quad \partial_\mu \varphi \sim j_\mu \quad \epsilon^\mu \partial_\nu \varphi \sim j'_\mu$$

The naive engineering dimension is $[\partial_\mu \varphi]_\ell = 1$. This shows us that $\partial_\mu \varphi$ does indeed have meaningful correlation functions.

The more interesting composite operators are the previously mentioned exponentials. So we choose the two simplest ones:

$$(1.13) \quad V_\pm = e^{\pm i\varphi}$$

These are known in the math and string theory literature as “vertex operators.” These show up all over various operator algebras in mathematics. Now we are interesting in

$$(1.14) \quad \langle V_+(x_1) \cdots V_+(x_n) V_-(y_1) \cdots V_-(y_m) \rangle$$

So what would make this meaningful? Can we take $m_0 \rightarrow 0$? In the presence of the IR regulator, we will also need a UV regulator, which we will call Λ . We want to trade this

^{1,4} This theorem is known by this name by high energy physicists, or by some, the Coleman-Mermin-Wagner theorem. There is also pressure from condensed matter physicists to call it the CHMW theorem (H for Hohenberg).

UV regulator for some renormalization group scale, called $\mu_{\text{RG}} \equiv \mu$. So the two things we want to do are send $m_0 \rightarrow 0$ and send $\Lambda \rightarrow \mu$. These will both make sense if the correlation functions are meaningful. So how do we do this? We will be able to calculate these correlation functions because they simplify to Gaussian integrals. Let us first simplify to the following case:

$$(1.15) \quad \langle V_+(x_1) V_-(x_1) \rangle = \left\langle \exp \left(i \int d^2x J(x) \varphi(x) \right) \right\rangle_S$$

$$(1.16) \quad = \exp \left(-\frac{1}{2} \int d^2x d^2y J(x) G(x-y, m_0) J(y) \right)$$

So $V_+(x_1)$ is a special case of

$$(1.17) \quad \exp \left(i \int d^2x J(x) \varphi(x) \right)$$

when $J = \delta^2(x_1) - \delta^2(x_2)$. So we can plug this in to get

$$(1.18) \quad \langle V_+(x_1) V_-(x_1) \rangle = \exp \left(-\frac{1}{2} 2G(0, m_0^2) + \frac{1}{2} 2G(x_1 - x_2, m_0^2) \right)$$

But we have an expression for G from (1.11). Now if we send $m_0 \rightarrow 0$ and we get a UV divergence. Now to regulate this, we have to introduce a cutoff. The simplest way is to put the system on a two-dimensional lattice with spacing a . So call $a = 1/\Lambda$ where Λ has the dimension of energy. So now whenever there is an $m_0 |0|$ replace with $m_0 a = m_0/\Lambda$. This could be more sophisticated, but we take this more pragmatic route. So now

$$(1.19) \quad \langle V_+(x_1) V_-(x_1) \rangle = \exp \left(\frac{g^2}{2\pi} \left(\log \frac{m_0}{\Lambda} - \log m_0 |x_1 - x_2| \right) \right)$$

$$(1.20) \quad = \left(\frac{1}{|x_1 - x_2| \Lambda} \right)^{g^2/(2\pi)}$$

This is somewhat nice since the IR regulators drop out, but the bad news is that this UV regulator doesn't. But in reality, this is a good thing, because we want a good result in terms of the renormalization scale μ_{RG} . Now the classical dimensions are $[\varphi]_\ell = 0$ $[e^{i\alpha\varphi}]_\ell = 0$ which develops an anomalous dimension γ . Recall the $\beta(g)$ function determines the effects of the coupling constants, and γ determines the effect given by difference in anomalous dimension. In any case, Λ is still in there, but in a very nice place, where we can introduce a simple treatment given by the introduction of μ_{RG} . This is the scale at which we measure the system. The operators that are being measured, are rescaled versions of these operators. So we should call these V_{+0} or V_{+B} to get bare or renormalized operators, but we won't overdo notation. Naively we get

$$(1.21) \quad \langle V_+(x_1) V_-(x_1) \rangle = \left(\frac{1}{\mu |x_1 - x_2|} \right)^{g^2/(2\pi)} \left(\frac{\mu}{\Lambda} \right)^{g^2/(2\pi)}$$

So we propose that

$$(1.22) \quad \sqrt{Z} V_\pm^R = V_\pm$$

where

$$(1.23) \quad \sqrt{Z} = \left(\frac{\mu}{\Lambda} \right)^{g^2/(2\pi)}$$

so now

$$(1.24) \quad \langle V_+^R V_-^R \rangle = \left(\frac{1}{\mu |x_1 - x_2|} \right)^{g^2/(2\pi)} \quad [V_\pm^R] = g^2/4\pi$$

even though $[V_\pm] = 0$. So we get a non-trivial scaling dimension as expected. Now recalling $[\partial_\mu \varphi] = 1$, we now have a list of observable operators: $\partial_\mu \varphi, V_\pm$.

Now we've made our renormalization, so this better work for the extended calculation. First recall:

$$(1.25) \quad \langle V_+(x) V_-(0) \rangle = \exp(G(1/\Lambda) - G(|x|))$$

Now we can finally calculate:

$$(1.26) \quad \langle V_+(x_1) \cdots V_+(x_n) V_-(y_1) \cdots V_-(y_m) \rangle = G_{\text{bare}}^{(n,m)}(x_1, \dots, y_m)$$

$$(1.27) \quad = m_0^{\frac{(n-m)^2 g^2}{2\pi}} \left(\frac{1}{\Lambda} \right)^{g^2 \frac{n+m}{4\pi}} \left(\frac{\prod_{i \neq j}^n |x_i - x_j| \prod_{i' \neq j'}^m |y_{i'} - y_{j'}|}{\prod_{i,j} |x_i - y_j|} \right)^{g^2/(2\pi)}$$

where we got the factor $n - m + n(n - 1) + m(m + 1) - 2nm = (n - m)^2$ from

$$(1.28) \quad J(x) = \sum_{i=1}^n \delta(x_i) - \sum_{j=1}^m \delta(y_j)$$

So this gives us that for $n \neq m$, we get $G^{(n,m)} \equiv 0$. Now this expression should suggest we have some symmetry. We know what this symmetry is. It is a global $U(1)$ continuous symmetry, where $\varphi(x) \rightarrow \varphi(x) + c$ where c is a spacetime independent constant. Even independent of the Goldstone considerations, we would have found this symmetry from this correlation function calculation. So now $n = m$ is forced on us, so we just need to calculate:

$$(1.29) \quad G_{(R)}^{n,n}(\dots) = \left(\frac{\prod_{i \neq j}^n \mu |x_i - x_j| \prod_{i \neq j}^n \mu |y_i - y_j|}{\prod_{i,j} \mu |x_i - y_j|} \right)^{g^2/(2\pi)}$$

Now the RG equation^{1.5} can be written:

$$(1.30) \quad 0 = \mu \frac{d}{d\mu} \left\langle \underbrace{V_+ \cdots V_i}_{2n} \right\rangle = Z^n \left(\mu \frac{\partial}{\partial \mu} + 2n\mu \frac{\partial \log \sqrt{Z}}{\partial \mu} \left\langle V_+^{(R)} \cdots V_-^{(R)} \right\rangle \right)$$

$$(1.31) \quad \implies \left(\mu \frac{\partial}{\partial \mu} + 2n\gamma \left(\frac{g^2}{4\pi} \right) \right) G_{(R)}^{(n,n)}$$

Now we can compare this to the standard RG equation:

$$(1.32) \quad \left(\mu \frac{\partial}{\partial \mu} + \beta(g_R) \frac{\partial}{\partial g_R} + n\gamma_+(g_R) + n\gamma_-(g_R) \right) G_{(R)}^{n,n}(x, y; g_R; \mu)$$

where β is the beta function, and $\gamma_\pm(g_R)$ are the anomalous dimensions of V_\pm . Notice $\beta(g_R) = 0$ for this theory. When $\beta(g_R)$ then we are sitting on the line of exact CFTs. Systems where you have families of exact CFTs are important for various reasons. There is a full class of CFTs of this type, with $c = 1$ where c is the central charge. We will return to this, since it shows off these boson fermion equivalences we are trying to derive. Note that the scalar theory has a $U(1)$ symmetry, but in the massless case there are even more. These will be easier to find in comparison with the fermion case.

^{1.5} This is also a Callan-Symanzik equation.

3. Fermionic field theory

We now consider free fermions in Euclidean, Wick rotated $\mathbb{R}^{1,1}$ Lorentz spacetime. This theory can be treated in a significantly simpler manner. We choose a Dirac fermion ψ (i.e. complex valued). We can compare this to having two Majorana fermions rather than one dirac fermion. On an infinite plane, these look the same, but if we are working on a more complicated underlying spacetime, then these things become very different when we take on the periodicity of non-contractible cycles. We start with an action of the form

$$(1.33) \quad S(\psi, \bar{\psi}) = \int d^2x \bar{\psi} \not{\partial} \psi$$

where $\bar{\psi}$ is the dirac conjugate. Now recall we have

$$(1.34) \quad \psi = \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix}$$

In the component notation, the action can be written

$$(1.35) \quad S = \int dz d\bar{z} (\bar{\psi}_+ \partial_{\bar{z}} \psi_+ + \bar{\psi}_- \partial_z \psi_-)$$

Notice we have a classical $U(1) \times U(1)$ symmetry, where

$$(1.36) \quad \psi \rightarrow e^{i\theta + i\theta_5 \gamma_5} \psi$$

$$(1.37) \quad \bar{\psi} \rightarrow e^{-i\theta + i\theta_5 \gamma_5} \bar{\psi}$$

So this already predicts, that if this is supposedly equivalent to a free scalar, then we better have another copy of $U(1)$ symmetry laying around in that theory.

The conserved current for each one of these $U(1)$ symmetries is

$$(1.38) \quad J_\mu = \bar{\psi} \gamma_\mu \psi$$

$$(1.39) \quad J_\mu^5 = i \gamma_5 \gamma_\mu \psi$$

Now writing this in terms of the components, we get

$$(1.40) \quad J_1 = \bar{\psi} \gamma_1 \psi = \bar{\psi}_+ \psi_+ + \bar{\psi}_- \psi_-$$

$$(1.41) \quad J_2 = i (\bar{\psi}_+ \psi_+ - \bar{\psi}_- \psi_-)$$

Now since we have the identity: $i \gamma_5 \gamma_\mu = -\epsilon_{\mu\nu} \gamma^\nu$, we see that these two conventions are related to one another from the ϵ symbol: $J_\mu^5 = -\epsilon_{\mu\nu} J^\nu$. So now we have the theory and its symmetries, and we want to calculate some correlation functions. In particular, we need the propagator:

$$(1.42) \quad \Delta_+ := \langle \bar{\psi}_+(z, \bar{z}) \psi_+(0) \rangle = -\frac{1}{2\pi z}$$

$$(1.43) \quad \Delta_- := \langle \bar{\psi}_-(z, \bar{z}) \psi_-(0) \rangle = -\frac{1}{2\pi \bar{z}}$$

And this is all we really need to calculate correlation functions. For example,

$$(1.44) \quad \left\langle \prod_1^n \bar{\psi}_+(z_i) \psi_+(z'_i) \right\rangle = \left(-\frac{1}{2\pi}\right)^n \det\left(\frac{1}{z_i - z'_i}\right)$$

$$(1.45) \quad \left\langle \prod_1^n \bar{\psi}_-(\bar{z}_i) \psi_-(\bar{z}'_i) \right\rangle = \left(-\frac{1}{2\pi}\right)^n \det\left(\frac{1}{\bar{z}_i - \bar{z}'_i}\right)$$

4. Equivalence

Now we move on to the equivalence of these two theories. This really means showing that the endpoint functions are the same. To do this, we need to massage the functions calculated in (1.45). We have the following identity to help us:

$$(1.46) \quad (-1)^{n+1} \det \left(\frac{1}{z_i - z'_j} \right) = \frac{\prod_{i < j} (z_i - z_j) \prod_{i' < j'} (z'_{i'} - z'_{j'})}{\prod_{i,j} (z_i - z'_j)}$$

PROOF. Proceed by induction. For $n = 2$, we can calculate:

$$(1.47) \quad \det \left(\frac{1}{x_i - y_j} \right) = \frac{1}{(x_1 - y_1)(x_2 - y_2)} - \frac{1}{(x_1 - y_2)(x_2 - y_1)}$$

$$(1.48) \quad = \frac{(x_1 - y_2)(x_2 - y_1) - (x_1 - y_2)(x_1 - y_2)}{\prod_{i,j} (x_i - y_j)}$$

and expand. □

Now we introduce

$$(1.49) \quad \sigma_+ = \bar{\psi}_- \psi_+ \sigma_- = \bar{\psi}_+ \psi_-$$

So we have

$$(1.50) \quad \left\langle \prod_1^n \sigma_+(x_i) \sigma_-(y_i) \right\rangle = (-1)^n \left\langle \prod \bar{\psi}_+(y_i) \psi_+(x_i) \right\rangle \left\langle \prod \bar{\psi}_-(x_i) \psi_-(y_i) \right\rangle$$

$$(1.51) \quad = \left(\frac{1}{2\pi} \right)^{2n} \left(\frac{\prod_{i < j} |x_i - x_j|^2 |y_i - y_j|^2}{\prod_{i,j} |x_i - y_j|^2} \right)^{g^2/(4\pi)}$$

so when $g^2 = 4\pi$, this exponent is just one. Now we get the equivalence between the correlation function from the bosonic theory:

$$(1.52) \quad \left(\frac{1}{2\pi} \right) \left\langle \prod_{i=1}^n \mu^2 V_+^{(R)}(x_i) V_-^{(R)}(y_i) \right\rangle$$

So our bosonization expression is

$$(1.53) \quad 2\pi \sigma_{\pm} = \mu V_{\pm}^{(R)}$$

Then with the expression

$$(1.54) \quad V_{\pm}^{(R)} = \left(\frac{\Lambda}{\mu} \right)^{g^2/(4\pi)} V_{\pm}$$

we get

$$(1.55) \quad \bar{\psi}_- \psi_+ = \frac{\Lambda}{2\pi} e^{i\varphi} \quad \bar{\psi}_+ \psi_- = \frac{\Lambda}{2\pi} e^{-i\varphi}$$

Now we notice that the Lagrangians formally match as well, so we can finish off our “dictionary:”

$$(1.56) \quad \bar{\psi} \not{\partial} \psi \leftrightarrow \frac{1}{2} (\partial_\mu \varphi)^2 \quad j^\mu = \bar{\psi} \gamma^\mu \psi \leftrightarrow \frac{\epsilon^{\mu\nu}}{\sqrt{\pi}} \partial_\nu \varphi$$

$$(1.57) \quad \bar{\psi} \psi \leftrightarrow -\Lambda \cos \sqrt{4\pi} \varphi \quad \bar{\psi} \gamma_* \psi \leftrightarrow \Lambda \sin \sqrt{4\pi} \varphi$$

Now call the current corresponding to the scalar U(1) shift symmetry: $\mathcal{J}^\mu = \partial_\mu \varphi$. Notice we aren't getting the same one here, we're getting a dual version:

$$(1.58) \quad \frac{\epsilon^{\mu\nu}}{\sqrt{\pi}} \partial_\nu \varphi \equiv j^\mu \sim \epsilon^{\mu\nu} \mathcal{J}_\nu$$

Later on we will call this an electric-magnetic duality. At this point, we might consider proposing another duality: the scalar field φ should also be equivalent to another scalar, say, θ , related by

$$(1.59) \quad \partial_\mu \varphi \equiv \epsilon_{\mu\nu} \partial^\mu \theta$$

So we can write this in terms of φ or θ . This is known as T-duality in the string theory literature. This seems strange, but in some sense, this is quite clear in 1+1 dimensions.

Let us now return to a Minkowski signature. We introduce light-cone coordinates: $x^\pm = t \pm s$. Then the KG equation is given by

$$(1.60) \quad 0 = \partial_+ \partial_- \varphi$$

$$(1.61) \quad \varphi = \psi_+(x^+) + \varphi_-(x^-)$$

so we get a left and right moving piece. Now we can just define:

$$(1.62) \quad \theta = \varphi_+(x^+) - \varphi_-(x^-)$$

so in either case, we are describing the same theory with two different scalar fields. It turns out this type of duality symmetry related systems at large and small R on the RG scale. This duality symmetry takes $R \rightarrow \bullet/R$. This is a refinement of the moduli space of these scalar fields in 1+1 dimensions with one free scalar field component. So our claim is that the dirac theory is somewhere on this scale at R_D , which is something like 4π . Then if we seek the location of the pairs of Majorana fermions, there is another radius further up, which has a branch moving off of this exact CFT line, which corresponds to some scalar field theory of $\tilde{\varphi}$ which takes values in $S^1/\mathbb{Z}_2 \equiv I$ which is known as an orbifold. Now, amazingly, there is a point on this other branch, which we call R_{majorana} which is equivalent to a pair of majorana fermions. So we are fillig in this space of theories with one scalar degree of freedom. So how do we evaluate how many degrees of freedom we have in some particular theory, if the fields aren't quantum objects? This is related to central charge and conformal anomaly, which we will be studying later. But there is a good set of correlation functions in these dimensions, and there is a good invariant, the central charge, which gives the number of DOF in a unitary CFT. Amazingly enough, all such theories have been classified for $c \leq 1$.

5. Massless Schwinger model

We now consider the application of bosonization to the massless Schwinger model. The action for this theory takes the form:

$$(1.63) \quad S_{\text{QED}} = \int d^2x \left\{ \bar{\psi} \not{\partial} \psi + A_\mu j^\mu - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \right\}$$

where j^μ is the electromagnetic current associated to the Dirac field. Note that when dealing with non-abelian Yang-Mills theories one often writes $1/(4e^2)$ to replace the coefficient of $(F^{\mu\nu})^2$ which is just amounts to rescaling the field. Recall $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. Note that in $d+1$ dimensions, we have $d-1$ polarizations. So here, we have 0 classical polarizations. But as we have seen, we should not trust classical Lagrangians. We might not actually have a propagating fermion and 0 polarizations. But even with 0 polarizations, this isn't trivial, because we might still have non-zero electric fields. So we might have a long-distance

field, even if we don't have propagating photons, and we will actually see, using the tool of bosonization that we even get a degree of freedom with non-zero mass.

Now trade these fermions $\psi, \bar{\psi}$ for φ using bosonization. Now write $A_\mu = \partial_\mu \alpha$ so when we make a scalar transformation, α just becomes $\alpha + \epsilon_{\mu\nu} \partial^\mu f(x, t)$. Now f is gauge invariant, so we can just locally go to a gauge where $\alpha = 0$, and essentially forget about α . Then this becomes $A_\mu = \epsilon_{\mu\nu} \partial^\nu f(x, t)$. Now we have some trivial identities in Minkowski signature:

$$(1.64) \quad \epsilon_{01} = +1 \quad \epsilon_{\mu\nu} \epsilon^{\nu\sigma} = \delta_\mu^\sigma$$

$$(1.65) \quad \epsilon^{01} = -1 \quad \eta_{\mu\nu} = \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix}$$

Then according to our discussion above, our action becomes:

$$(1.66) \quad S = \int d^2x \left\{ \frac{1}{2} (\partial_\mu \varphi)^2 + \frac{\partial_\mu \varphi}{\sqrt{\pi}} \partial^\mu f + \frac{1}{2e^2} \partial^2 f \partial^2 f \right\}$$

where $\partial^2 = \partial^\mu \partial_\mu$. So our first result, is that this complicated fermionic theory is really just a simple bosonic theory. So instead of working to see we don't have any special fermionic effects, we can just use bosonization to treat it as a Gaussian theory, where we can readily make exact calculations. Note our source in the Gaussian integration is $J = -\partial^2 f / \sqrt{\pi}$. So we can write:

$$(1.67) \quad S = \int d^2x \left\{ -\frac{1}{2\pi} \partial_\mu f \partial^\mu f + \frac{1}{2e^2} \partial^2 f \partial^2 f \right\}$$

$$(1.68) \quad = \int d^2x \left\{ \frac{1}{2\pi} A_\mu A^\mu + \frac{1}{2e^2} A_\mu \partial^2 A^\mu \right\}$$

Now note that in momentum space, we get a dispersion relation: $\omega^2 = k^2 + e^2/\pi$ so we do have one propagating degree of freedom, with a mass. Originally Schwinger was trying to make the point that gauge invariance does not imply a massless theory. Now we can calculate the two-point functions in momentum space:

$$(1.69) \quad \langle A_\mu A_\nu \rangle = \left(\eta_{\mu\nu} - \frac{p_\mu p_\nu}{p^2} \right) \frac{1}{p^2 - e^2/\pi}$$

Note this is not a free field approximation, this is the exact two point function. This has no cut, only a pole at the mass of the physical particle.

DEFINITION 1.1. A field theory is said to exhibit *confinement*, when there are no asymptotic fermionic states with non-zero charge.

In other words, we have shown confinement in this theory.

We can also show that this theory exhibits spontaneous breaking of the chiral symmetry. This is very similar to what we expect from QCD. We can argue this by evaluating:

$$(1.70) \quad \langle \sigma_{pm} \rangle = \frac{\Lambda}{2\pi} \langle e^{+i\theta} \rangle = \exp \left(-\frac{1}{2} \left(-\frac{1}{2\pi} \log \frac{m}{\Lambda} + \dots \right) 4\pi \right) = \frac{m}{2\pi} e^{\dots}$$

again using the relations from bosonization. Note the mass of this scalar θ is $e/\sqrt{2\pi}$. So all the Λ terms cancel in the last step.

EXERCISE 5.1. Check that at short distances ($x \rightarrow 0$) we have that:

$$(1.71) \quad \langle \sigma_+(x) \sigma_-(0) \rangle \sim \frac{1}{2\pi^2 x^2}$$

6. Thirring model - sin-Gordon theory

We will develop a duality between two theories, neither of which are free using the following correspondences:

$$(1.72) \quad \bar{\psi} \not{\partial} \psi \longleftrightarrow \frac{1}{2} (\partial_\mu \varphi)^2$$

$$(1.73) \quad \bar{\psi} \gamma^\mu \psi \longleftrightarrow \frac{\epsilon^{\mu\nu}}{\sqrt{\pi}} \partial_\nu \varphi$$

$$(1.74) \quad \bar{\psi} \psi \longleftrightarrow -\Lambda \cos(\sqrt{4\pi} \varphi)$$

$$(1.75) \quad \bar{\psi} \gamma_* \psi \longleftrightarrow \Lambda \sin(\sqrt{4\pi} \varphi)$$

$$(1.76) \quad (\bar{\psi} \psi)^2 \longleftrightarrow \left[-\Lambda \cos(\sqrt{4\pi} \varphi) \right]^2 \approx -\frac{1}{2\pi} (\partial_\mu \varphi)^2$$

Recall that

$$(1.77) \quad \langle \sigma_\pm \rangle \sim \Lambda e^{G(0, m^2)} \sim m$$

So the UV cutoff showing up here is important because it shows us this is not some silly technical triviality, and Λ plays a critical role in a full understanding of these theories.

6.1. Sin-Gordon theory. The most practical way to discover the sine-Gordon theory, would be to start with some fermionic theory, and ask what theory it is dual to. It was however actually discovered on its own, and later seen to be dual to a fermionic theory.

The action for this theory is:

$$(1.78) \quad S_{\text{SG}} = \int d^2x \left\{ \frac{1}{2} \partial_\mu \varphi \partial^\mu \varphi + \frac{m^4}{\lambda} \left\{ \cos\left(\frac{\sqrt{\lambda}}{m} \varphi\right) - 1 \right\} \right\}$$

we just subtract 1 such that for $\varphi = 0$ we get 0 energy. Note that we have introduced two couplings. The coupling m will play the role of mass, and λ will play the role of self coupling. Now notice for small values of φ , we get that the first term is roughly:

$$(1.79) \quad \approx \frac{1}{2} (\partial_\mu \varphi)^2 - \frac{1}{2} m^2 \varphi^2 + \frac{\lambda}{4!} \varphi^4 - \frac{\lambda^2}{m^2} \frac{\varphi^6}{6!}$$

We will not prove it here, but we offer the fact:

Proposition 1.2. *The sine-Gordon theory is classically integrable.*

Note integrability has many different definitions. The practical definition is a theory which can be rewritten completely in terms of action-angle variables. That is, taking variables $p, q \rightarrow I, \theta$ where I is some action. Then the Hamiltonian becomes very simple. Then a theory possesses quantum integrability when we can solve for the S -matrix exactly.

We now state some elementary facts about this theory: The equation of motion does indeed contain sin:

$$(1.80) \quad \partial^\mu \partial_\mu \varphi + \frac{m^3}{\sqrt{\lambda}} \sin\left(\frac{\sqrt{\lambda}}{m} \varphi\right) = 0$$

We can gather from this, that the theory also has an infinite number of degenerate ground states. We can always ask for perturbations around such minima, and then using perturbation theory to get some Gaussian integrals and such. But we can also have theories which jump from one to another. Well $\varphi = 0$ is a solution, but now we want to impose a boundary condition for which φ is at one minimum at $-\infty$ in space, and in the neighboring minimum

at ∞ in space. To get this, we introduce a static^{1.6} “soliton” solution. A soliton is a generic name for a quantum object which does not dissipate.

We can anticipate that this will be 0 at $-\infty$ and then jump to the next level between there and ∞ on plot of $\bar{\varphi}$ versus \bar{x} where we have introduced:

$$(1.81) \quad \bar{x} = mx \quad \bar{t} = mt \quad \bar{\varphi} = \frac{\sqrt{\lambda}}{m} \varphi$$

Now we note some symmetries of the theory.

$$(1.82) \quad \bar{\varphi} \rightarrow -\bar{\varphi} \quad \bar{\varphi} \rightarrow \bar{\varphi} + 2\pi n, n \in \mathbb{Z}$$

We can now generalize and impose the boundary condition that $\bar{\varphi}(-\infty) = 2\pi N_1$ is the minimum of the potential for integer N_1 and $\bar{\varphi}(\infty) = 2\pi N_2$ for integer N_2 . Then define $Q = N_1 - N_2$ to be the topological quantum number. Now we can actually express:

$$(1.83) \quad Q = \frac{1}{2\pi} \int d\bar{x} \partial_{\bar{x}} \bar{\varphi} \sim \int d\bar{x} j^0 \quad j^\mu = \epsilon^{\mu\nu} \partial_\nu \bar{\varphi}$$

Note this current is conserved because $\partial_\mu (\epsilon^{\mu\nu} \partial_\nu \bar{\varphi}) = 0$ identically. Note that $\varphi \rightarrow \varphi + c$ with current $\tilde{j}^\mu = \partial^\mu$ is not a symmetry here.

Now we can actually write down a solution. The simplest solution has $Q = \pm 1$. We can check trivially that

$$(1.84) \quad \bar{\varphi}(x) = \pm 4 \arctan(e^{-\bar{x} - \bar{x}_0})$$

is a solution. The theory has many more, such as an analytically known soliton anti-soliton solution. This solution corresponds to scattering of a soliton off of an anti-soliton, and is therefore a time dependent solution.

$$(1.85) \quad 4 \arctan \left\{ \frac{\sinh \left(\frac{u\bar{t}}{\sqrt{1-u^2}} \right)}{u \cosh \left(\bar{x} / \sqrt{1-u^2} \right)} \right\}$$

We also have a “breather solution” given by

$$(1.86) \quad \bar{\varphi} = 4 \arctan \left\{ \frac{\sin(v\bar{t} / \sqrt{1+v^2})}{v \cosh(\bar{x} / \sqrt{1+v^2})} \right\}$$

This one is more of a bound state. It is time dependent in the sense that it oscillates out and back, but represents one compact object. It is sometimes called a doublet.

We might wonder why this Soliton solution is considered to be non-perturbative. One way to see this is to notice that we can evaluate the energy classically by plugging this into the classical energy formula, which gives

$$(1.87) \quad E_{\wedge \text{soliton}, \ell} = \frac{8m^3}{\lambda}$$

Clearly this is non-perturbative, since turning λ off doesn't make any sense. This inverse scaling with couplings is very typical for these solitonic objects. We can still handle them in this particular example since this theory is integrable.^{1.7} Quantum mechanically, the energy can be calculated as:

$$(1.88) \quad E_{\text{soliton}, qm} = \frac{8m^3}{\lambda} + \frac{m}{\lambda} + \mathcal{O}(\lambda)$$

^{1.6} We know how to boost static solutions to arbitrary velocity, so this actually gives us a whole family of soliton solutions.

^{1.7} One often requires supersymmetry in a non-integrable theory to control our solitons.

Now the sine-Gordon model, in the notation $\varphi' = \frac{\sqrt{\lambda}}{m}\varphi$, can be rewritten

$$(1.89) \quad S_{\text{SG}} = \frac{m^2}{\lambda} \int \frac{1}{2} (\partial_\mu \varphi')^2 + m^2 \cos \varphi'$$

the important thing to notice here, is that the coupling constant $\kappa^2 = \frac{\lambda}{m^2}$ is present with a power of -2 . Some of the literature takes $t = \kappa^2$.

6.2. Thirring model and duality. The Thirring model Lagrangian is written

$$(1.90) \quad S_T = \int d^2x \left\{ \bar{\psi} (\not{\partial} + m_F) \psi - \frac{1}{2} g J_\mu J^\mu \right\}$$

Now we can apply the duality sequence to various terms in this action:

$$(1.91) \quad S = \int d^2x \left\{ (\partial_\mu \varphi)^2 + \Lambda \cos \varphi' + \frac{g}{\pi} (\partial_\mu \varphi)^2 \right\}$$

Then to make contact with the sine-Gordon model, we see that the appropriate equation is:

$$(1.92) \quad 1 + \frac{g}{\pi} = \frac{4\pi}{\kappa^2}$$

The main implication of this is that $m_F = 0$.

But where does the fermion show up in the sine-Gordon theory? Mandelstam showed that the fermion is exactly the soliton in the sine-Gordon theory. If we decompose the scalar into the left chiral and left chiral component, we had

$$(1.93) \quad \sigma_\pm \bar{\psi} \psi \sim e^{\# \varphi} = e^{\# \varphi_+} e^{\# \varphi_-}$$

which suggests the above.

7. Theory classification

Recall the classification space of all CFTs in $1 + 1$ dimensions which have the same “number of degrees of freedom” as a real scalar. The horizontal axis is parameterized by the radius R of S^1 . We have $R = 0$ on the far left, and some critical radius $R = R_*$ where the theories with $0 < R < R_*$ are in correspondence with the theories with $R > R_*$ where we decompose

$$(1.94) \quad \varphi = \varphi_+(x^+) + \varphi_-(x^-) \theta = \varphi_+(x^+) - \varphi_-(x^-)$$

$$(1.95) \quad R \rightarrow \frac{\#}{R} \quad \int (\partial \varphi)^2 \rightarrow \int (\partial \theta)^2$$

and send The same correspondence holds for the vertical axis parameterized by R_I , where R_I is the radius of $S^1/(\mathbb{Z}/2\mathbb{Z})$. Amazingly enough, these two one dimensional families of free field theories, touch each other at a certain radius, where $R_I = R_{I*}$. Note this is not where $R = R_*$. The crossing point corresponds to $R_I = R_{I*}$ and $R > R_*$.

We have proven that the Dirac fermion sits somewhere between $R = R_*$ and the point where the R_I axis crosses the R axis. we call this R_D . As it turns out, the pair of Majorana fermions (which would be equivalent on local spacetime to a Dirac fermion) live somewhere on the R_I axis. Explicitly, the Majorana² theory lives at $R_I = R_M$.

Also recall the theory at $R = R_D$ had a $U(1) \times U(1)$ symmetry. In fact, this symmetry is present along the whole R axis. Additionally, if we consider $R = R_*$, this symmetry gets enhanced and becomes an $SU(2) \times SU(2)$ symmetry. So when we have the $(U(1))^2$ symmetries, there will be currents $\partial_\mu \varphi$, $\epsilon^{\mu\nu} \partial_\mu \varphi$, each with dimension one. Now for the

enhanced case at $R = R_*$, some of the allowed $e^{\pm i\varphi_{\pm}}$ will also get dimension 1, so they can also be candidate conserved currents.

EXERCISE 7.1. Show that these candidate currents do indeed satisfy the proper commutation relations.

In addition to the theories which are parameterized by R and R_I , we have three isolated theories living in the bottom right. These are given by the symmetries of the dodecahedron, the octahedron, and the icosahedron. We call these D, O , and I . The reason we mention this, is because this classification of CFTs, turns out to be one of many mathematical classification problems known as an ‘‘ADE’’ classification. This means there are three sequences of objects. The A sequence of objects, $\{A_n\}_{n \in \mathbb{N}}$, the D sequence of objects, $\{D_n\}_{n \in \mathbb{N}}$, and the E sequence $\{E_6, E_7, E_8\}$. In this problem the R axis gives the A sequence, the R_I axis gives the D sequence, and the E sequence is given by these D, O , and I theories. As it turns out, these show up because of the $SU(2)$ symmetry. As we saw, the appropriate vertex operators conspire to get this new symmetry, so there is an underlying $SU(2)$ which we can start with. Then we can ask the following interesting question: what are all of the possible discrete subgroups $\Gamma \subseteq SU(2)$? For each such Γ , we can then consider $SU(2)/\Gamma$. Note this is precisely mapped to an ADE classification problem. There is an infinite A sequence of subgroups which contain \mathbb{Z}_n an infinite D sequence which contain $\mathbb{Z}_2 \times \mathbb{Z}_n$ and then E_6, E_7 , and E_8 . Now factoring out by an A_n puts us on R axis, factoring out by some D_n puts us on the R_I axis, and then the E s give you these three loose points.

So applying what we learned in the sin-Gordon Thirring duality, we see that we actually have extended the duality to the whole line using self-interacting massless theories. Similarly we could show the same for the Majorana theories along this vertical axis.

Non-abelian Yang-Mills

Recall that QED is a $U(1)$ gauge theory with field A_μ . Non-abelian Yang-Mills theories are essentially an attempt to generalize this notion to general Gauge fields to describe other interactions, such as the strong and weak interactions, with some A_μ^I where I is indexed by the adjoint representation of some Lie group G . This will typically be a compact Lie group since this will give us positive energy and a well defined Hamiltonian and Lagrangian. The issue is of course that other interactions have different non-Abelian symmetries. Unless explicitly mentioned, we will consider relativistic Minkowski space-time with Poincare symmetries.

Yang Mills is a QFT of a “connection” described by A_μ^a which is the adjoint representation of some compact Lie group G . We can more compactly write this is $A_\mu^a T^a$ which we will see more about soon.

1. Lie groups

DEFINITION 2.1. The group $U(N)$ consists of unitary $n \times n$ matrices.

This just means that $U^\dagger = U^{-1}$.

Proposition 2.1. *A Lie group also forms a manifold.*

DEFINITION 2.2. The Lie algebra associated to a Lie group G is the tangent space at the identity: $T_e G = \mathfrak{g}$.

DEFINITION 2.3. The adjoint representation of a Lie group is the action of group on its own algebra.

REMARK 2.1. Physicists are sloppy about differentiating between Lie algebras and Lie groups. It is often easiest to figure out which one is being talked about from the context. For example if someone is taking a tensor product \otimes of two objects, they’re likely dealing with groups. If they’re taking the sum \oplus , they’re likely dealing with algebras.

Unless stated otherwise, we will be working only with compact Lie groups.

DEFINITION 2.4. A subgroup $H \subseteq G$ is a subgroup such that H is preserved by the conjugate group action of G on H .

DEFINITION 2.5. A Lie group is *simple* iff it is connected as a topological manifold, non-abelian,^{2.1} and does not have a non-trivial connected normal subgroup.

DEFINITION 2.6. A Lie group is *semi-simple* iff it is a product of any finite number of simple Lie groups.

EXAMPLE 2.1. $SU(2)$ connected and non-abelian, and $U(1) \subseteq SU(2)$, but this subgroup is not normal. Therefore $SU(2)$ is simple. $SU(3)$ is as well. $U(1)$ is not simple.

^{2.1} Ruling out $U(1)$

EXAMPLE 2.2. The standard model group $U(1) \otimes SU(2) \otimes SU(3)$ is not semi-simple because of the $U(1)$ factor.

1.1. Killing-Cartan classification. This classification of Lie groups will give us the available choices of Yang-Mills gauge theories. In particular, we will have a classification given A_n, B_n, C_n, D_n , for $n \in \mathbb{Z}^+$, and then the special cases $E_{6,7,8}, F_4, G_2$. This is an extended version of the ADE classification we saw before.

For any Lie group G , there is a unique Lie algebra \mathfrak{g} . Now if we work with the Jacobi identities, this implies a lot of constraining structure, not only on \mathfrak{g} , but on a certain sub-algebra.

Vector spaces by themselves don't necessarily have a commutation relation. This would have to be extra data we provide. So why does \mathfrak{g} already have this? Well if we take any vector v in \mathfrak{g} , then we can use the group action to "parallel transport" this vector v to any point in the group. Now we can define $v \rightarrow V_L$ which is a left invariant vector field. In particular, we can define the Lie commutation relation for the Lie algebra with $[X(g), Y(g)]$.

We now introduce the Cartan sub-algebra. For any Lie algebra we have this commutation relation. So we might ask whether there is some maximum commuting sub-algebra? In other words the largest set of elements:

$$(2.1) \quad \{t_i \in \{T^a\} \mid [t_i, t_j] = 0\}$$

This set, if it exists, is unique up to conjugation. In any case, for any \mathfrak{g} , there is indeed such a sub-algebra \mathfrak{h} . This will correspond to some subgroup of G , which will correspond to some particular generators. Note that this subgroup will always be Abelian by construction. Now we can consider some coordinate system $e^{i\theta^a T^a}$, and if we consider only those corresponding to \mathfrak{h} , $e^{i\theta^i t^i}$, then the subgroup generated by this must be $U(1)^{\text{rank}(G)}$ to some power called the rank of G .

Now since G is a finite dimensional manifold, it of course has a dimension, and now we have defined the rank as well. For $SU(2)$, the number of such generators is just 1. Here $U(1)$ corresponds to the Cartan sub-algebra. We also have $U(1)^2 \subseteq SU(3)$. This notion of rank is important, because it gives us our classification.

One we have the Cartan subgroup^{2.2} we go back to the cartan subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$, and call $n = \text{rank}(G)$. We now have an n -dimensional real subspace inside the real algebra. We also know when we exponentiate into the full group, this is a compactification of the full manifold into a Torus. This means we need a discrete lattice $\Gamma \subseteq \mathbb{R}^n$ which will have n basis elements called roots, and there will be some geometric structure in place. The generic idea is that the structure of this lattice completely determines the groups, and therefore we can classify them accordingly. We sketch a proof.

SKETCH. As it turns out, up to overall normalization, for all Lie groups, there are only three possible "lengths" of the roots. We call these l_1, l_2, l_3 .

Let's start producing elements of a graphical representation. We will associate a vertex with each root, and denote its interaction with the other roots by the path we draw between them. Recall for rank n we want to associate n roots.

It also turns out to be the case, that only two of these three lengths can show up and be present for any particular Lie group.

There are two conventions we can take now. We can either write nodes corresponding to roots of shorter length as \circ , and write the longer length roots as \bullet , or we can denote the longer root with an arrow pointing along the path between them, towards the longer root.

^{2.2} Also called the Cartan "torus," since S^1 is the topology of $U(1)$.

We take the second conventions. We also write paths between l_1 nodes as a single line. We write paths between l_1 and l_2 as double lines, and we write paths between l_1 and l_3 as triple lines. These diagrams can be seen in fig. 1. \square

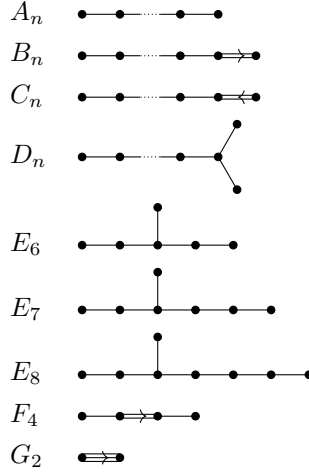


FIGURE 1. Dynkin diagrams

These seem somewhat esoteric and useless, besides maybe giving us a way to classify these group. But in reality, these can show us quite a bit. In particular, we can find what are often called “accidental isomorphisms” by comparing Dynkin diagrams:

$$(2.2) \quad B_2 = C_2 \quad O(5) = \text{Sp}(2)$$

$$(2.3) \quad D_3 = A_3 \quad O(6) = \text{SU}(4)$$

Also note the following dimensions $\dim E_n = 78, 133, 248$, $\dim F_4 = 52$, $\dim G_2 = 14$

Recall the ADE classification is a special case of this. This is because these cases are what are called simply laced. This just means that all the nodes are representing roots of the same “length” so there are only single lines.

2. Generalizing scalar fields

We understand scalar field theories relatively well. We also know we can promote this to a complex scalar field, which can of course carry a charge. So now take φ to be a complex field. Now in some sense we promote the

$$(2.4) \quad \partial_\mu \varphi \rightarrow D_\mu(A) \varphi (\partial_\mu - ieA_\mu) \varphi$$

When we replace with this covariant derivative, we not only make the theory invariant, but we also make it an interacting theory.

Now we can easily replace φ with an n dimensional scalar field $\{\varphi^i(x)\}_{i=1}^N$ which has a global $U(N)$ symmetry. We could replace this with any lie group G . Any compact lie group can be used for Yang-Mills quantization. This is because we need a positive definite

metric. This metric must be invariant under the group action, and it's a fact compact lie groups do this. We write this group action as:

$$(2.5) \quad U : \varphi(x) \rightarrow U\varphi(x)$$

now we postulate that

$$(2.6) \quad S = \int \partial\bar{\varphi}\partial\varphi \pm m^2(\bar{\varphi}\varphi) - \frac{\lambda}{4!}(\bar{\varphi}\varphi)^2$$

Now this is clearly a nice scalar field theory, which is invariant under a global symmetry. The notion of a global symmetry here, means a symmetry which acts regardless of the space-time location. These global symmetries can be broken, which means the vacuum state in the field theory is not preserved by the charges which generate the field theory: $Q|0\rangle \neq 0$. If it is not broken, then this leads to an actual physical symmetry acting on states.

Now we can ask if we can gauge this global $U(N)$ symmetry. In other words, can we replace this with a symmetry which is allowed to depend on the space-time location. So this lie group now becomes a space-time dependent function which takes values in $U(n)$.^{2,3} This is what is referred to as a local symmetry, or gauge symmetry. Recall in the case of the scalar field, we had to introduce the gauge field at this step. We now attempt to do the same:

$$(2.7) \quad \partial_\mu\varphi \rightarrow \partial_\mu(U\varphi) = U\partial_\mu\varphi + (\partial_\mu U)\varphi - U[\partial_\mu\varphi + (U^\dagger\partial_\mu U)]\varphi$$

So now by brute force we have shown

$$(2.8) \quad D_\mu\varphi^i = \partial_\mu\varphi^i - iA_{\mu j}^i(x)\varphi^j(x)$$

where A_μ is valued in the adjoint representation of the group G .

Now we have a Lie group of symmetries, where $U \in G = U(N)$. So in some small neighborhood of e , we can write

$$(2.9) \quad U(x) = e^{i\theta^a(x)T_a}$$

where we have used a as the index of the adjoint rep of G , and θ is a local coordinate. So for $\dim G = n$, there will be a basis in the n dimensional Lie algebra space, and $a \in \{1, \dots, n\}$. Often times people just write both a indices as covariant indices, since when G is a compact Lie group^{2,4} there is a unique invariant killing metric, g_{ab} . There is a theorem which says this is not only positive definite, but also diagonalizable, so $g_{ab} = \pm\delta_{ab}$. So θ is a set of n coordinate on some patch in the neighborhood of e .

So what is the required transformation rule for A_μ ? We can guess

$$(2.10) \quad A_\mu \rightarrow A_\mu + i\theta^a[T^a, A_\mu] + \partial_\mu\theta^a(x)T^a$$

which is a generalization of the abelian case for QED.

WARNING 2.1. Note we have hidden the charge e . So where is the coupling constant here? We have implicitly rescaled A_μ to A'_μ to absorb whatever gauge coupling shows up. This will show up in the kinetic term in the Gauge fields instead. We choose between it showing up here, and showing up in the covariant derivative.

We now remind ourselves that $[A_\mu] = [\partial_\mu] = 1$ since the terms in our definition must have the same dimension. When we get to Yang mills, the only dimensionless coupling is in 4 dimensions, so this is important.

^{2,3} This means it becomes a principal section of a bundle.

^{2,4} This will always be the case for us.

Another important convention distinction is that in physics: $A_\mu^P = A_\mu$. However mathematicians instead define: $A_\mu^M = -iA_\mu^P$. Therefore a mathematician's covariant derivative is:

$$(2.11) \quad D_\mu = (\partial_\mu + A_\mu^M)$$

whereas a physicist has a $-i$ in there. This is a reflection of the general trend that physicists like unitary matrices to be similar to e^{ir} where r is real. Then T^a just has to be hermitian. In the mathematician convention, this just makes T^a anti-hermitian. Then the commutation relations become

$$(2.12) \quad [T^a, T^b] = if^{ab}_c T^c$$

where the f are the structure coefficients which are real. So the idea is, to eliminate this i , don't introduce it. In the physics convention, the ugly i appears in the commutation relation, but not in

$$(2.13) \quad A_\mu^a \rightarrow A_\mu^a + \partial_\mu \theta^a - f^{abc} \theta^b A_\mu^c$$

REMARK 2.2. Note that exchanging a, b in the f s is antisymmetric, and a priori c is unrelated. But when we get the metric structure, we get this antisymmetric behavior in all three indices.^{2.5} In particular, compact groups automatically have a positive definite metric on their lie algebra, and when we raise and lower indices on compact groups, then f_{abc} are completely antisymmetric, for any simple compact lie group.

Note that physicists call A_μ^a a "gauge field" and mathematicians call it a "component representation of a connection." The idea here, is that what the physicist interprets as a funny one form, comes from a very natural mathematical construction: Consider some manifold M . For us, this will just be \mathbb{R}^3 .¹ Then for any compact G , define the principal bundle P to be a manifold somehow "bigger" than M , equipped with a projection where the inverse under this projection is called a fiber of this principal bundle. So far this is any fibration, but a principal bundle is when each fiber $F_x \cong G$. So we essentially have a copy of G above every point. So for every x there is some neighborhood of x , U_x , such that when we look at the stack of things projecting into this neighborhood, this is of the form $U_x \times G$. So G acts on this big manifold P by right multiplication, which moves us around the give fiber. Then a space-time dependent gauge transformation would be a section Γ of P . So a connection tells us how to parallel transport. Now if the values are in a lie group, then we want to transfer a lie algebra from one location to another.

Note that Gauge symmetries are effectively just redundancies in our description. We are not saying that $A_\mu \rightarrow \tilde{A}_\mu$ are two states, we are saying they are two different mathematical descriptions of the same state. So when interested in correlation functions which are invariant under gauge transformations, it is important to maintain this viewpoint.

Recall in QED, $F_{\mu\nu}$ was gauge invariant. Showing the analogous result here is a bit harder. First of all we will have to define $F_{\mu\nu}$, but this will only be gauge covariant, not gauge invariant. This just means

$$(2.14) \quad \sum_a F_{\mu\nu}^a F_{\sigma\rho}^a$$

might be invariant. Until otherwise stated, we take the convention $A_\mu \equiv A_\mu^M = -iA_\mu^P$. It is useful to think of $A = A_\mu dx^\mu$ as some sort of locally matrix valued one form. Indeed for each choice of connection, there is a one form called ω whose coordinate transformation on

^{2.5} See the appendix in [23] for details.

space-time boils down to those of A_μ . So this is really just a slight abuse of notation. Now finally we get the transformation law:

$$(2.15) \quad A \rightarrow UAU^\dagger + U dU^\dagger$$

$$(2.16) \quad A_\mu^a \rightarrow A_\mu^a + \partial_\mu \theta(x)^a - f^{abc} \theta(x)^b A_\mu^c$$

where $d : \Lambda^p M \rightarrow \Lambda^{p+1} M$ and $d^2 = 0$.

Now with this property, we can check:

$$(2.17) \quad dA \rightarrow U dAU^\dagger + dU \wedge AU^\dagger - UA \wedge dU^\dagger + dU \wedge dU^\dagger$$

which is complicated. So we should really define $A^a \wedge A^b f^{abc}$: $[A \wedge A]$ to find that

$$(2.18) \quad F = dA + A \wedge A$$

which transforms as covariantly as possible:

$$(2.19) \quad F \rightarrow UFU^\dagger$$

Another interesting fact, is if we take $D = d + A$, then we can show, that $D^2 = F$. So in the two opposing notational systems we get:

$$(2.20) \quad F_{\mu\nu}^{\text{math}} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu] \quad F_{\mu\nu}^{\text{physics}} = \partial_\mu A_\nu - \partial_\nu A_\mu - i[A_\mu, A_\nu]$$

In any case, we can now write down the YM action:

$$(2.21) \quad S_{\text{ym}} = -\frac{1}{2g_{\text{ym}}^2} \int d^4x \operatorname{tr}(F_{\mu\nu} F^{\mu\nu})$$

where $\operatorname{tr} F^2 \sim g_{ab} F_{\mu\nu}^a F_{\mu\nu}^b$ is unique up to a normalization. This will of course affect what the YM coupling constant will be. In particular, we have

$$(2.22) \quad \operatorname{tr}(T^a T^b) = \# \delta^{ab}$$

where $\#$ is determined by about a million different potential conventions.

3. Quantization

3.1. Initial attempts. In this section, following Faddeev-Popov in [10], we settle on the convention:

$$(2.23) \quad \operatorname{tr}(T^a T^b) = 2\delta^{ab}$$

which will result in a factor of $1/8$ in the action.^{2,6}

We start with the same route as we did in QED. Recall the trick was finding by inspection, that

$$(2.24) \quad \langle A_\mu A_\nu \rangle \sim \frac{\eta_{\mu\nu}}{p^2 + i\epsilon}$$

then this worked if the matter described by this field had a conserved current.

More formally, we want to integrate

$$(2.25) \quad \int \mathcal{D}A_\mu \exp\left(\frac{i}{8g^2} \int \operatorname{tr}(F_{\mu\nu} F^{\mu\nu})\right)$$

Now this of course diverges badly. So we need to figure out how to quantize systems of this sort. Recall that a lesson we learned earlier, is that path integrals are a good shorthand for quantizing a system, but only if we already know how to quantize in the Hamiltonian formalism.

^{2,6} One has to simply get used to all of the different conventions here, and just focus on consistency.

Recall that we worked to get the following (naive) compact expression:

$$(2.26) \quad \mathcal{Z} = \int \mathcal{D}q(t) \exp\left(i \int dt \mathcal{L}(q, q', t)\right)$$

but this sometimes required corrections. But the Hamiltonian formalism gave us the (more) correct expression:

$$(2.27) \quad \mathcal{Z} = \int \mathcal{D}p(t) \mathcal{D}q(t) \exp\left(i \int dt (pq' - H)\right)$$

which sometimes reduces to the same in the coordinate language, but not always.

REMARK 2.3. It is a beautiful and fundamental statement in general, that there is a canonical measure in phase space. In particular, we can take the exterior derivative:

$$(2.28) \quad d(p_i dq^i) = dp_i \wedge dq^i = \omega$$

where ω is a symplectic form, and ω^n is a volume form.

So now we fix the following:

$$(2.29) \quad \mathcal{A}_\mu := A_\mu^a T^a A_\mu \rightarrow \mathcal{A}_\mu^\omega := \omega(x) \mathcal{A}_\mu \omega^{-1}(x) + \partial_\mu \omega(x) \omega^{-1}(x)$$

$$(2.30) \quad S = \frac{1}{8g^2} \int \text{tr}(F_{\mu\nu} F_{\mu\nu}) \quad F_{\mu\nu} = \partial_\nu A_\mu - \partial_\mu A_\nu + [A_\mu, A_\nu]$$

Then the equations of motion are:

$$(2.31) \quad D^\mu F_{\mu\nu} = 0$$

Note that this is where we notice there is a non-invertibility of the Gaussian term because we see we don't have linearly independent equations of motion. Formally, this is because $D^\mu D^\nu F_{\mu\nu} \equiv 0$.

People like Feynman^{2.7} wanted to postulate that the propagator should still just be

$$(2.32) \quad D_{\mu\nu} = \frac{\eta_{\mu\nu}}{p^2 + i\epsilon}$$

After all this worked fine in QED, because we could just ignore the non-invertible portion of the propagator.

We now make our way towards the Hamiltonian formalism. We postulate that the Lagrangian takes the form:

$$(2.33) \quad \mathcal{L} = E_k^a \partial_0 A_k^a - h(E_k, A_k) + A_0^a C^a$$

Inspired by the analogy $E_k^a, A_k^b \sim p_i q^i$, and

$$(2.34) \quad h = \frac{1}{2} \left\{ (E)^2 + (G)^2 \right\}$$

Specifically, we rewrite the trace:

$$(2.35) \quad \text{tr} \left\{ \left(\partial_\nu \mathcal{A}_\mu - \partial_\mu \mathcal{A}_\nu + g [A_\mu, A_\nu] - \frac{1}{2} \mathcal{F}_{\mu\nu} \right) \right\}$$

Up to a total derivative, this will become a Hamiltonian form:

$$(2.36) \quad -\frac{1}{2} \text{tr} \left\{ E_k \partial_0 A_k - \frac{1}{2} (E_j^2 + G_k^2) + \mathcal{A}_0 \mathcal{C} \right\}$$

^{2.7}In the early 60s, Feynman was attempting to treat Yang-Mills quantization as he had treated QED, and was shocked by the non-unitarity of the S -matrix under the tricks which worked for QED.

where

$$(2.37) \quad E_k = \mathcal{F}_{k0} G_k = \frac{1}{2} \epsilon^{ijk} \mathcal{F}_{ji} h = \frac{1}{2} \left\{ (E_k^a)^2 + (G_k^a)^2 \right\}$$

Note that \mathcal{C} is a Lie algebra valued object such that

$$(2.38) \quad \mathcal{C} = \partial_k E_k - g[\mathcal{A}_k, E_k] \quad \mathcal{C} = C^a T^a$$

this h will be called a Hamiltonian even though is not a proper Hamiltonian as we are used to. Specifically, this is a degenerate Lagrangian, and therefore doesn't lead to an actual Hamiltonian. Therefore we need to apply Dirac's more sophisticated Hamiltonian treatment.

We have the canonical pair E, A so we get Poisson brackets:

$$(2.39) \quad \{E_k^a(x^i), E_l^b(x^j)\}_{\text{PB}} = \delta_{kl} \delta^{ab} \delta(x^i - x^j)$$

$$(2.40) \quad \{C^a(x^i), C^b(y^j)\}_{\text{PB}} = g f^{abc} C^c(x) \delta(x^i - y^j)$$

Our system is a generalized Hamiltonian system in the dense of Dirac. We now try to figure out how to treat them.

3.2. Dirac's formalism. Consider a general system with a finite number of degrees of freedom. Then there is a classical system Γ with n degrees of freedom, and phase space Γ^{2n} of dimension $2n$ with coordinates p_i, q_i for $i \in \{1, \dots, n\}$. The equation of motion is $\varphi^\alpha(p, q) = 0$, and we postulate that the action of this system is

$$(2.41) \quad S = \int \left\{ \sum_{i=1}^n p_i q'_i - h(p, q) - \sum \lambda^\alpha \varphi^\alpha(p, q) \right\} dt$$

where $\alpha \in \{1, \dots, m\}$ for $m < n$. We now assume two things. We already know p, q satisfy canonical Poisson brackets. We also have

$$(2.42) \quad \{h, \varphi^\alpha\}_{\text{PB}} = C^{\alpha\beta}(p, q) \varphi^\beta$$

$$(2.43) \quad \{\varphi^\alpha, \varphi^\beta\}_{\text{PB}} = C^{\alpha\beta\gamma}(p, q) \varphi^\gamma$$

We will such a system, a system with first class constraints. It is somewhat easy to see that Yang Mills is such a theory.

First class constraints are:

$$(2.44) \quad \{h, \varphi^\alpha\} = C^{\alpha\beta}_{(p,q)} \varphi^\beta \quad \{\varphi^\alpha, \varphi^\beta\} = \sum C^{\alpha\beta\gamma}_{(p,q)} \varphi^\gamma$$

Second class constraints are

$$(2.45) \quad \{\varphi^\alpha, \varphi^\beta\} \sim \epsilon^{\alpha\beta}$$

We will be focusing on first class constraints. For more information on the general theory, see [12].

CLAIM 2.1. The system Γ , is equivalent to a simpler looking system Γ^* with $n - m$ canonical pairs (degrees of freedom).

The phase space $\Gamma^{*2(n-m)}$ can be constructed as follows.

- (1) In addition to the data we already have, pick m additional conditions $\chi^\alpha(p, q) = 0$. These can be anything as long as:

$$(2.46) \quad \det \{\varphi^\alpha, \varphi^\beta\} \neq 0 \quad \{\chi^\alpha, \chi^\beta\} = 0$$

We will see why these conditions are useful as we develop the reduced phase space.

(2) Identify $2m$ conditions, which is that we force

$$(2.47) \quad \chi^\alpha = \varphi^\alpha = 0$$

This gives $\Gamma^{*2(n-m)}$.

(3) Find canonical coordinates p^*, q^* . We will call this an “adapted coordinate system.” We will have

$$(2.48) \quad q = (\chi^\alpha, q^*) \quad p = (p^\alpha, p^*)$$

Now in these variables, if we look at the first condition in the last step, we see that as a consequence of this, we have

$$(2.49) \quad \det \left| \frac{\partial \varphi^*}{\partial p^\beta} \right| \neq 0$$

In principal, we can now invert $\varphi^\alpha(p, q) = 0$ to solve for $p^\alpha = p^\alpha(p^*, q^*, \chi^\alpha)$. So this is not an independent momentum anymore. We now do this manually. So we have defined $\Gamma^{*2(n-m)}$ such that $\chi^\alpha = \varphi^\alpha = 0$. So this means p^*, q^* are a good set of canonical coordinates, since $p^\alpha = p^\alpha(p^*, q^*)$. So the hamiltonian:

$$(2.50) \quad h^*(p^*, q^*) = h(p, q)|_{\varphi=\chi=0}$$

So now we just need to calculate \mathcal{Z} with the measure replaced with our new canonical coordinates. We now prove that the original Γ with constraints is the same system as Γ without:

PROOF. We know the equation of motion for Γ is:

$$(2.51) \quad p'_i + \frac{\partial h}{\partial q_i} + \lambda^\alpha \frac{\partial \varphi^\alpha}{\partial q_i} = 0$$

$$(2.52) \quad q'_i + \frac{\partial h}{\partial p_i} - \lambda^\alpha \frac{\partial \varphi^\alpha}{\partial p_i} = 0$$

and $\varphi^\alpha = 0$. So now solving for $\lambda^\alpha(t)$, we note that λ^α are arbitrary, but imposing $\chi^\alpha = 0$ fixes this. Of course in our usual language, χ^α is a gauge fixing condition. Now the standard hamiltonian of Γ^* is:

$$(2.53) \quad q'^* = \frac{\partial h^*}{\partial p^*} \quad p'^* = \frac{\partial h^*}{\partial q^*}$$

now go to the adapted system with Γ . Specifically focus on the q^α equation of motion to get

$$(2.54) \quad \frac{\partial h}{\partial p_\alpha} + \lambda^\alpha \frac{\partial \varphi^\beta}{\partial p^\alpha} = 0$$

Now it is important to notice that (2.47) allows us to express λ^α uniquely. Now moving onto the equations of motion for the remaining q^* , we get that in Γ :

$$(2.55) \quad q'^* = \frac{\partial h}{\partial p^*} + \lambda^\alpha \frac{\partial \varphi^\alpha}{\partial p^*} = 0$$

And in Γ^* :

$$(2.56) \quad q'^* = \frac{\partial h^*}{\partial p^*} = \frac{\partial h}{\partial p^*} + \frac{\partial h}{\partial p^\alpha} \frac{\partial h}{\partial p^*}$$

now when we set these equal, we get some condition where we have dropped relative signs, to get:

$$(2.57) \quad \lambda^\alpha \left(\frac{\partial \varphi_\alpha \varphi}{\partial p^*} + \frac{\partial \varphi_\alpha}{\partial \varphi_\beta} \frac{\partial \varphi_\beta}{\partial \varphi^*} \right)$$

so we are done for q , and we leave the p case as an exercise. \square

We could apply this directly to YM, but let's continue in this more familiar language of QM. For Γ , we define the path integral:

$$(2.58) \quad \mathcal{Z} = \int \mathcal{D}p^*(t) \mathcal{D}q^*(t) \exp \left(i \int p^* q'^* - h^*(p^*, q^*) \right)$$

now we know this is perfectly accurate, by the preceding discussion, so the statement is:

Proposition 2.2.

$$(2.59) \quad \mathcal{Z} = \int \mathcal{D}p(t) \mathcal{D}q(t) \mathcal{D}\lambda(t) \prod_{t,\alpha} \delta(\chi^\alpha) \prod_t \det |\{\varphi_\alpha, \chi_\beta\}|$$

$$(2.60) \quad \exp \left(i \int p q' - h(p, q) - \lambda^\alpha \varphi^\alpha(p, q) \right)$$

The determinant is called the Faddeev-Popov determinant.

PROOF. First we integrate over λ which imposes the constraint conditions. So now we are on the constraint surface $\varphi = 0$. So this gives us something looking somewhat like

$$(2.61) \quad \int \mathcal{D}p \mathcal{D}q \exp \left(i \int p q' - h \right) \prod \delta(\varphi^\alpha) \delta(\chi^\alpha) \prod \det |\{\varphi^\alpha, \chi_\beta\}|$$

now go to the coordinates $(\chi^\alpha, q^*), (p(p^*, q^*), p^*)$. Recall that on this constraint surface,

$$(2.62) \quad \prod \delta(\varphi^\alpha) \delta(\chi^\alpha) \prod \det |\{\varphi^\alpha, \chi_\beta\}| = \prod \delta(\varphi_\alpha) \delta(q_\alpha) \det \left| \frac{\partial \varphi_\alpha}{\partial p_\beta} \right|$$

$$(2.63) \quad = \prod \delta(q_\alpha) \delta[p_\alpha - p_\alpha(p^*, q^*)]$$

Now by inspection, we can integrate out p_α and q_α , to get the expected expression. \square

3.3. YM as a theory with second class constraints. We can now apply this picture to the actual YM theory. Recall that

$$(2.64) \quad \{E_k^a(x^i), A_l^b(y^j)\} = \delta_{kl} \delta^{ab} (x^i - y^j)$$

$$(2.65) \quad \{C^a(x^i), X^b(y^j)\} = g f^{abc} C^c(x^i) \delta(x^i - y^j)$$

This means we can calculate:

$$(2.66) \quad \left\{ \int d^3x \left[(E_k^a)^2 + (G_k^a)^2 \right], C^b(y^i) \right\} = 0$$

so this is indeed a first class set of constraints.

So now we do step one, and fix the Coulomb gauge $\partial_k A_k = 0$.

WARNING 2.2. It is tempting to set the Gauge fixing condition to be the Lorentz gauge $\partial^\mu A_\mu = 0$ but this does not only depend on p, q , it also depends on the Lagrange multiplier. So this is out for now.

This also satisfies the conditions of the second step, because clearly

$$(2.67) \quad \{\partial_k \mathcal{A}_k, \partial_i \mathcal{A}_i\} = 0$$

Now we have to calculate:

$$(2.68) \quad \left\{ C^a(x^i), \partial_k A_k^{(b)}(y) \right\} = -\partial_k [\partial_k \delta^{ab} - g f^{abc} A_k^c(x)] \delta(x-y) = M_c[\mathcal{A}]$$

So we can develop a perturbative expansion in g . The leading term here will be $\pm \partial_k \partial_k \delta^{ab}$, so this is an invertible matrix. Note that we don't have Lorentz invariance at this point, but we still have the full expression:

$$(2.69) \quad \mathcal{Z} = \int \mathcal{D}E_i^a(x) \mathcal{D}A_\mu^a(x) \exp\left(i \frac{1}{8g_{\text{ym}}^2} \int d^4x \text{tr}(F_{\mu\nu} F^{\mu\nu})\right) \prod \delta(\partial_k \mathcal{A}_k) \prod \det M_c[\mathcal{A}] .$$

WARNING 2.3. Still treat E_i and A_μ as independent variables. Recall, that luckily, the E_i factors only appear in (up to) quadratic terms.

EXERCISE 3.1. Take the full Hamiltonian formula, and see that plugging in Yang-Mills gives us this expression.

This M_c matrix is an operator (since we're talking about field theory after all). In particular,

$$(2.70) \quad M_c = \Delta \delta^{ab} - g f^{abc} A_k^c \partial_k$$

which is the analogue of what would just be $\partial_\mu \partial^\mu$ is a scalar φ^4 theory. Importantly, this expression depends on A . We have two conclusions from this:

- (1) QED works without fully understanding the nature of this determinant. This is just a lucky coincidence. As we will see, to treat the FP determinant in full generality requires a fair bit more work.
- (2) If we want to avoid viewing this pesky determinant, there is a clear path to avoiding it completely. In particular, we can come up with some gauge fixing, such that the matrix is independent of the Gauge field A .

4. Gauge fixing in YM

The general point of this section, is to find how we can move from the Coulomb gauge fixing $\partial_i \mathcal{A}_i = 0$ to the Lorentz gauge fixing $\partial_\mu \mathcal{A}^\mu = 0$, since we obviously want our theory to be Lorentz invariant.

To understand the relationship between the theory and the Gauge fixing we choose, it is convenient to introduce an object Δ_L called the Faddeev-Popov (FP) determinant. This is some object associated to a given Gauge fixing condition L , such that:

$$(2.71) \quad \Delta_L(\mathcal{A}) \int \prod_x \delta(\partial_\mu \mathcal{A}_\mu^\omega) d\omega = 1$$

Note that we write $\partial_\mu \mathcal{A}_\mu \equiv \partial_\mu \mathcal{A}^\mu$ in order to reserve the upper index location to denote more features. In this case we use it to denote the the group action. So here, \mathcal{A}_μ^ω is the gauge transformation of \mathcal{A}_μ under a local group element ω :

$$(2.72) \quad \mathcal{A}_\mu^\omega = \omega^{-1} \mathcal{A}_\mu \omega + \omega^{-1} \partial_\mu \omega$$

In the above definition of Δ_L , we take $d\omega$ to be any G invariant measure on G . This is known as the Haar measure. In particular, we just define it to be whatever measure satisfies:

$$(2.73) \quad d\omega\omega_0 = d\omega = d\omega_0\omega$$

So this resolution of the identity is given by an integral over the group G , and spacetime.

Now we have the following statement about Δ_L :

Proposition 2.3. $\Delta_L(\mathcal{A})$ is gauge invariant.

PROOF. We can write:

$$(2.74) \quad \Delta_L(\mathcal{A}^{\omega'})^{-1} = \int \prod \delta(\partial_\mu \mathcal{A}_\mu^{\omega'\omega}) d\omega$$

$$(2.75) \quad = \prod \delta(\partial_\mu \mathcal{A}_\mu^{\tilde{\omega}}) d(\omega'^{-1}\tilde{\omega})$$

$$(2.76) \quad = \prod \delta(\partial_\mu \mathcal{A}_\mu^{\tilde{\omega}}) d(\tilde{\omega})$$

$$(2.77) \quad = \Delta_L(\mathcal{A})^{-1}$$

where we have taken $\tilde{\omega} := \omega'\omega$. □

We also have the following:

Proposition 2.4. On a a gfc surface,

$$(2.78) \quad \prod \det M_c[\mathcal{A}] = \Delta_c(\mathcal{A})$$

PROOF. When $\partial_k \mathcal{A}_k = 0$, we have $\omega = e$. So in order to perform this integral, we only need to concentrate on the vicinity of e . In other words, we can perform this only over elements of the Lie algebra. So write $\omega \approx e + u(x)$ for u some small deviation. So we need to evaluate $\partial_k \mathcal{A}_k^\omega$ for such an ω . But this is just

$$(2.79) \quad \partial_k \mathcal{A}_k^\omega = \Delta u - g[\mathcal{A}_k, \partial_k u] = M_c[\mathcal{A}]$$

Now evaluating this integral gives us $\Delta_X(\mathcal{A}) = \det(M_c(\mathcal{A}))$ so we are done. □

Now taking the resolution of the identity from the definition of Δ_L , we can rewrite the path integral as:

$$(2.80) \quad \mathcal{Z} = \int \mathcal{D}\mathcal{A}_\mu(z) \Delta_L(\mathcal{A}) \int \prod \delta(\partial_\mu \mathcal{A}_\mu^\omega) d\omega$$

$$(2.81) \quad \exp\left(\frac{i}{8g_{\text{ym}}^2} \int d^4x \text{tr}(F_{\mu\nu}F^{\mu\nu})\right) \prod \delta(\partial_k \mathcal{A}_k) \Delta_c(\mathcal{A})$$

$$(2.82) \quad \rightarrow \int \mathcal{D}\mathcal{A}_\mu(z) \Delta_L(\mathcal{A}) \int \prod \delta(\partial_\mu \mathcal{A}_\mu) d\omega$$

$$(2.83) \quad \exp\left(\frac{i}{8g_{\text{ym}}^2} \int d^4x \text{tr}(F_{\mu\nu}F^{\mu\nu})\right) \prod \delta(\partial_k \mathcal{A}_k^{\omega^{-1}}) \Delta_c(\mathcal{A})$$

$$(2.84) \quad = \int \mathcal{D}_\mu(x) \Delta_L \prod \delta(\partial_\mu \mathcal{A}^\mu) \exp\left(\frac{i}{8g_{\text{ym}}^2} \int d^4x \text{tr}(F_{\mu\nu}F^{\mu\nu})\right)$$

where we brought $\mathcal{A}_\mu \rightarrow \mathcal{A}_\mu^{\omega^{-1}}$ in the second line.

We now have two more tricks:

- (1) The first trick will bring us one step closer to Feynman rules. We notice that we are falling short in two major ways. First of all, we can't really work with the FP determinant perturbatively yet. We also don't really want to work directly with the product of delta distributions in our path integral expression. In other words, we want to write this \mathcal{Z} as an integral of the form:

$$(2.85) \quad \mathcal{Z} = \int \mathcal{D}\mathcal{A}_\mu(x) \cdots \exp(iS_{\text{eff}}(\mathcal{A}, \cdots))$$

where we might need to introduce additional integration variables. So instead of $\partial^\mu \mathcal{A}_\mu^a = 0$, we take $\partial^\mu \mathcal{A}_\mu^a = \mathcal{B}^a(x)$ where \mathcal{B} is called an auxiliary field. This lets us rewrite:

$$(2.86) \quad \mathcal{Z} = \int \mathcal{D}_\mu(x) \mathcal{D}B^a(x) \Delta_L(\mathcal{A}) \prod \delta(\partial_\mu \mathcal{A}^\mu - \mathcal{B})$$

$$(2.87) \quad \exp\left(\frac{i}{8g_{\text{YM}}^2} \int d^4x \text{tr}(F^2)\right) \exp\left(i\chi \int \text{tr}(\mathcal{B}) d^4x\right)$$

Note that if we integrate over \mathcal{B} we get back to the delta function form.

- (2) The second trick is to notice that the determinant of any bosonic matrix like this, can be written like:

$$(2.88) \quad \det M_x(\mathcal{A}) = \int \mathcal{D}c(x) \mathcal{D}b(x) \exp\left(i \int d^4x (b(x) M_c[\mathcal{A}] c(x))\right)$$

where b, c are Grassmannian variables with the same field structure as the gauge fixing condition. The field c is called a *ghost* and the field b is an *anti-ghosts*. The YM notation often takes c as the ghost and \bar{c} as the anti-ghost. We sometimes write it this way.

4.1. Alternative gauge choices. The Gauge fixing we assumed above a priori lets us think the ghost anti-ghost don't decouple from the gauge field. Can we find gauges where we decouple b, c ?

For U(1) gauge theories, they were already decoupled. For non-abelian theories, we will get the wrong result from leaving off the ghost terms. For non-abelian theories there is a beautiful class of "axial" gauges. In general, they rely on a choice of a constant vector v^μ . This breaks Lorentz since we are choosing a direction in spacetime. Then the gauge condition is just $v^\mu A_\mu^a = 0$.

EXAMPLE 2.3. If we take $v^\mu = (1, 0, 0, 0)$, then $A_0^a = 0$ which is often called a "temporal" gauge.

We can also take v^μ such that $v^\mu v_\mu = 0$. This condition will be called "lightcone" gauge.^{2.8}

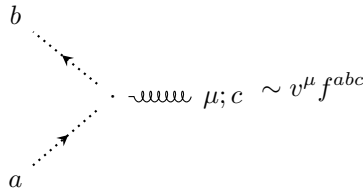
Generically, if v is spacelike, then this is referred to as "axial proper." People often pick the third direction, so they set $A_3^a = 0$ which simplifies Feynman rules.

Let's now derive the propagator in an axial gauge:

$$(2.89) \quad \Pi_{\text{axial}}^{ab\mu\nu} = \frac{1}{p^2 + i\epsilon} \left[-\eta^{\mu\nu} + \frac{v^\mu p^\nu + v^\nu p^\mu}{vp} - \frac{(v^2 + \xi p^2) p^\mu p^\nu}{(vp)^2} \right] \delta^{ab}$$

^{2.8} This should really be called infinite momentum gauge.

Now observe that this satisfies $p_\mu \Pi^{\mu\nu} = 0$ for $p^2 = 0$. In addition, if we take $\xi = 0$, then $v_\mu \Pi^{\mu\nu} = 0$. So

(2.90) 

so ghosts decouple, but in a more subtle way than in QED.

Alternative gauge choices: The Gauge fixing we assumed above a priori lets us think the ghost anti-ghost don't decouple from the gauge field. Can we find gauges where we decouple b, c ?

For U(1) gauge theories, they were already decoupled. For non-abelian theories, we will get the wrong result from leaving off the ghost terms. For non-abelian theories there is a beautiful class of “axial” gauges. In general, they rely on a choice of a constant vector v^μ . This breaks Lorentz since we are choosing a direction in spacetime. Then the gauge condition is just $v^\mu A_\mu^a = 0$.

EXAMPLE 2.4. If we take $v^\mu = (1, 0, 0, 0)$, then $A_0^a = 0$ which is often called a “temporal” gauge.

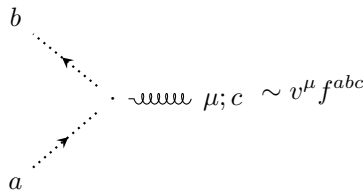
We can also take v^μ such that $v^\mu v_\mu = 0$. This condition will be called “lightcone” gauge.^{2.9}

Generically, if v is spacelike, then this is referred to as “axial proper.” People often pick the third direction, so they set $A_3^a = 0$ which simplifies Feynman rules.

Let's now derive the propagator in an axial gauge:

(2.91)
$$\Pi_{\text{axial}}^{ab\mu\nu} = \frac{1}{p^2 + i\epsilon} \left[-\eta^{\mu\nu} + \frac{v^\mu p^\nu + v^\nu p^\mu}{vp} - \frac{(v^2 + \xi p^2) p^\mu p^\nu}{(vp)^2} \right] \delta^{ab}$$

Now observe that this satisfies $p_\mu \Pi^{\mu\nu} = 0$ for $p^2 = 0$. In addition, if we take $\xi = 0$, then $v_\mu \Pi^{\mu\nu} = 0$. So

(2.92) 

so ghosts decouple, but in a more subtle way than in QED.

^{2.9} This should really be called infinite momentum gauge.

5. Feynman rules

We now use the preceding developments to attempt to formulate Feynman rules for our YM theory. Our most “Feynman diagram friendly” form of the path integral is:

$$(2.93) \quad \mathcal{Z} = \int \mathcal{D}A_\mu(x) \mathcal{D}b(x) \mathcal{D}c(x) \mathcal{D}B(x) \delta(\partial_\mu A^\mu - \mathcal{B})$$

$$(2.94) \quad \exp\left(i \int d^4x \left\{ \frac{1}{2} \text{tr}(F_{\mu\nu} F^{\mu\nu}) + \xi \text{tr}(\mathcal{B}^2) + ibM_c(\mathcal{A})c \right\}\right)$$

where one often defines:

$$(2.95) \quad S_{\text{ym}} = \text{tr}(F^{\mu\nu} F_{\mu\nu}) \quad S_{\text{gf}} = \xi \text{tr}(B^2)$$

Here where ξ is called a gauge fixing parameter. If we are ever trying to make some sort of physical prediction, it is always a good check to assure that this result has no ξ dependence, since no results should depend on any Gauge fixing condition.

Propagator for Gauge fields A_μ (“gluons”) is given by

$$(2.96) \quad \nu; b \xrightarrow{p} \mu; a = \frac{i}{p^2 + i\epsilon} (-\eta^{\mu\nu} + (1 - \xi) p^\mu p^\nu / p^2) \delta^{ab}$$

Notice that the gluon propagator is the photon propagator with this extra δ^{ab} factor.

The scalar propagator is given by:

$$(2.97) \quad j \xrightarrow{p} i = \frac{i\delta^{ij}}{p^2 - m^2 + i\epsilon}$$

The fermionic propagator is given by:

$$(2.98) \quad j \xrightarrow{p} i = \frac{i\delta^{ij}}{\not{p} - m + i\epsilon}$$

The propagator for the ghost fields b and c is given by:

$$(2.99) \quad \nu; b \xrightarrow{p} \mu; a = \frac{i\delta^{ab}}{p^2 + i\epsilon}$$

For the triple gluon vertex, the derivative can act on any of the gluons, so we have the Feynman rule:

$$(2.100) \quad \nu; b \xrightarrow{p} \cdot \begin{array}{l} \swarrow \text{uuuu} \mu; a \\ \searrow \text{uuuu} \rho; c \end{array} = gf^{abc} [g^{\mu\nu} (k-p)^\rho + g^{\nu\rho} (p-q)^\mu + g^{\rho\mu} (q-k)^\nu]$$

Note that since we take all the momentum incoming, $p + k + q = 0$.

The four-gluon vertex gives us:

$$(2.101) \quad \begin{array}{l} \mu; a \\ \rho; c \end{array} \cdot \begin{array}{l} \swarrow \text{uuuu} \nu; b \\ \searrow \text{uuuu} \sigma; d \end{array} = -ig^2 \times \left[\begin{array}{l} f^{abe} f^{cde} (g^{\mu\rho} g^{\nu\sigma} - g^{\mu\sigma} g^{\nu\rho}) \\ + f^{ace} f^{bde} (g^{\mu\nu} g^{\rho\sigma} - g^{\mu\sigma} g^{\nu\rho}) \\ + f^{ade} f^{bce} (g^{\mu\nu} g^{\rho\sigma} - g^{\mu\rho} g^{\nu\sigma}) \end{array} \right]$$

The following ghost vertex is given by the rule:

$$(2.102) \quad \begin{array}{c} \mu; b \\ \downarrow \text{wavy} \\ \begin{array}{ccc} & \cdot & \\ \nearrow \text{dotted} & & \searrow \text{dotted} \\ c^c & & b^a \end{array} \end{array} = -gf^{abc}p^\mu$$

The vertex for the Fermion interaction is given by:

$$(2.103) \quad \begin{array}{c} \mu; a \\ \downarrow \text{wavy} \\ \begin{array}{ccc} & \cdot & \\ \nearrow \text{solid} & & \searrow \text{solid} \\ j & & i \end{array} \end{array} = ig\gamma^\mu T_{ij}^a$$

Notice that this is the same as in the case of QED where the orientation of the vertex in the actual diagram does not matter.

Finally we have the scalar vertices:

$$(2.104) \quad \begin{array}{c} \mu; a \\ \downarrow \text{wavy} \\ \begin{array}{ccc} & \cdot & \\ \nearrow \text{dashed } k & & \searrow \text{dashed } q \\ j & & i \end{array} \end{array} = ig(k^\mu + q^\mu)T_{ij}^a$$

$$(2.105) \quad \begin{array}{c} \mu; b \\ \downarrow \text{wavy } m \\ \begin{array}{ccc} & \cdot & \\ \nearrow \text{dashed } k & & \searrow \text{dashed } q \\ j & & i \end{array} \\ \nu; a \\ \downarrow \text{wavy } u \end{array} = ig^2 T_{ik}^a T_{kj}^b g^{\mu\nu}$$

5.1. Representation theory. If we recall the loop corrections calculated in QED and QCD, it turns out that the more general YM corrections are the same up to some factors. We now lay some brief group theoretic formalism pertaining to these factors.

Define some object T , for a given representation, such that

$$(2.106) \quad \text{tr}(T_R^a T_R^b) = T(R) \delta^{ab}$$

Then this implies that

$$(2.107) \quad T(\text{fund}) = T_F = \frac{1}{2}$$

that is, $T_{ji}^a T_{ij}^b = 1/2 \delta^{ab}$. For the adjoint representation we have

$$(2.108) \quad T(\text{adj}) = T_A = N$$

that is, $f^{acd} f^{bcd} = N \delta^{ab}$.

The quadratic Casimir: Instead of taking the tr over the d dimensional space of the representation, we can take the product in some representation:

$$(2.109) \quad T_R^a T_R^a = C_2(R) \mathbf{1}_R$$

Then it is easy to check that the following is true:

$$(2.110) \quad \dim(R) C_2(R) = T(R) \dim(G)$$

We will call $C_2(F) = C_F = N^2 - 1 / (2N)$ and $C_2(A) = C_A = N$. These relations are used in effectively every QCD calculation.

5.2. One-loop corrections. We now want to calculate the β function to show QCD with $SU(N)$ is asymptotically free. So we want to calculate all of the one loop corrections we can think of. First the coulomb propagator, but also for

$$(2.111) \quad \dots \rightarrow \dots \cdot \begin{array}{l} \nearrow \\ \searrow \end{array}$$

Note we will be using the renormalized perturbation theory to set this calculation up. We have seen a good amount of this. We start by writing the bare, field, and write the Lagrangian using this bare field, bare mass, and the coupling as bare couplings. Then we introduce a renormalized field, a renormalized mass, and renormalized couplings. So we get a bunch of counterterms. So here we also do this. We denote these counterterms by

$$(2.112) \quad \text{---} \otimes \text{---}$$

We now apply the methods of dimensional regularisation. Commit to spacetime dimension $d = 4 - \epsilon$, and a renormalization group scale such that $[\mu] = 1$. It will turn out to be useful to also use $\tilde{\mu} := 4\pi e^{-\gamma_e} \mu$ where γ_e is the Euler constant. We now have:

$$(2.113) \quad S \sim -\frac{1}{2g_{\text{ym}}^2} \int \text{tr}(F_{\mu\nu} F^{\mu\nu}) d^d x$$

so we peel off the dimensional dependence, by taking the dimensionful $\tilde{g}_y m = \mu^{(4-d)/2} g_{\text{ym}}$ since g_{ym} is dimensionless. We know ξ should drop out in the end, but while making calculations, will write $\xi = 1$ and restore this in the end. The propagator is extremely simple:

$$(2.114) \quad \Pi^{ab\mu\nu} = \delta^{ab} \frac{-i\eta^{\mu\nu}}{p^2 + i\epsilon}$$

After these preliminaries, we can now start calculating diagrams. We calculate the effects of the fermionic loop first. As mentioned earlier, we are doing something very similar to the QED case:

$$(2.115) \quad i\mathcal{M}_F^{ab\mu\nu} = \text{---}^p \cdot \begin{array}{c} \curvearrowright^{k-p} \\ \curvearrowleft^p \end{array} \cdot \text{---}^p$$

$$(2.116) \quad = -\text{tr}[T^a T^b] (ig)^2 \int \frac{d^4 k}{(2\pi)^4} \frac{i}{(p-k)^2 - m^2} \frac{i}{k^2 - m^2}$$

$$(2.117) \quad \text{Tr}[\gamma^\mu (\not{k} - \not{p} + m) \gamma^\nu (\not{k} + m)]$$

In particular, we notice that this is exactly the QED case, only now we have the factor $\text{tr} [T^a T^b] = T_F \delta^{ab}$ to account for the colors. We will see

$$(2.118) \quad \mathcal{M} = -g^2 \delta^{ab} (\eta^{\mu\nu} p^2 - p^\mu p^\nu) \Pi_2(p^2)$$

which is required by gauge symmetry. So this is required by the fermion loop, and for the sum of the four loop gluon, the gluon bubble, and the ghost bubble, but not for these three diagrams on their own. We have the following expression:

$$(2.119) \quad M_F = \delta_{ab} T_F \left(\frac{g^2}{16\pi^2} \right) (\eta^{\mu\nu} p^2 - p^\mu p^\nu) \left[-\frac{8}{3} \frac{1}{\epsilon} - \frac{20}{9} - \frac{4}{3} \log \left(-\frac{\tilde{\mu}^2}{p^2} \right) \right]$$

so we have isolated our first one loop divergence. Now the sum for all four diagrams is given by:

$$(2.120) \quad \mathcal{M}^{ab\mu\nu} = \delta^{ab} \frac{g^2}{16\pi^2} (\dots)^\mu \times \left[C_A \left(\frac{10}{3\epsilon} + \frac{5}{3} \log \frac{\tilde{\mu}^2}{-p^2} \right) - n_f T_F \left(\frac{8}{3} \frac{1}{\epsilon} + \frac{4}{3} \log \left(\frac{\tilde{\mu}^2}{-\partial^2} \right) \right) \right]$$

The next step is to write the renormalized perturbation theory.

$$(2.121) \quad \mathcal{L} = -\frac{1}{4} (\partial_\mu A_\nu^B - \partial_\nu A_\mu^B)^2 + \dots$$

$$(2.122) \quad = -\frac{1}{4} Z_3 (\partial_\mu A_\nu^a - \partial_\nu A_\mu^a) + Z_2 \bar{\psi}_i (i\not{\partial} - Z_m m) \psi_i$$

$$(2.123) \quad -Z_{3c} \bar{c}^a \square c^a - g_R Z_{A^3} f^{abc} (\partial_\mu A_\nu^a) A_\mu^b A_\nu^c$$

$$(2.124) \quad -\frac{1}{4} g_R^2 Z_{A^4} (f^{eab} A_\mu^a A_\nu^b) (f^{ecd} A_\mu^c A_\nu^d)$$

$$(2.125) \quad + g_R Z_1 A_\mu^a \bar{\psi}_i \gamma^\mu T_{ij}^a \psi_j + g_R Z_{1c} f^{abc} (\partial_\mu \bar{c}^a) A_\mu^b c^c$$

where we have written $\varphi_0 \equiv \varphi_B$ for bare quantities, and φ_R for renormalized. Whenever a quantity has no subscript, it is assumed to be renormalized. Note that

$$(2.126) \quad Z_X = 1 + \delta_x$$

Now we have to calculate these Z_X in the perturbation theory. As we saw in the one loop calculation for the diagram:

$$(2.127) \quad \text{mmm} \otimes \text{mmm}$$

we get the expression:

$$(2.128) \quad \delta_3 = \frac{1}{\epsilon} \frac{g^2}{16\pi^2} \left[\frac{10}{3} C_A - \frac{8}{3} n_f T_F + (1 - \xi) C_A \right]$$

Now again from the previous calculations, we can calculate the counter-terms:

$$(2.129) \quad \delta_2 = \frac{1}{\epsilon} \frac{g^2}{16\pi^2} [-2C_F + 2(1 - \xi) C_F] \quad \delta_m = \frac{1}{\epsilon} \frac{g^2}{16\pi^2} (-6C_f)$$

Recall we have the following results:

$$(2.130) \quad \text{Diagram: a vertex with three external lines (two outgoing, one incoming)} = ig\Gamma_{ij}^{a\mu}$$

$$(2.131) \quad \text{Diagram: a vertex with three external lines (two outgoing, one incoming) and a loop (wavy line)} = ig(T^c T^a T^b)_{ij} \delta^{bc} \Gamma_{(2A)}^\mu$$

Now since we have the following expressions:

$$(2.132) \quad (T^b T^a T^b)_{ij} = (C_F - C_A/2) (T^a)_{ij}$$

$$(2.133) \quad \Gamma_{(2A)}^\mu = F_1^{(2A)} \left(\frac{P^2}{m^2} \right) \gamma^\mu + \frac{i\sigma^{\mu\nu}}{2m} p_\nu F_2^{(2A)} \left(\frac{p^2}{m^2} \right)$$

so we have the final expressions for the counter terms:

$$(2.134) \quad \delta_1 = \frac{1}{\epsilon} \left(\frac{g^2}{16\pi^2} \right) \left[-2C_F - 2C_a + 2(1-\xi)C_F + \frac{1}{2}(1-\xi C_A) \right]$$

$$(2.135) \quad \delta_2 = \frac{1}{\epsilon} \left(\frac{g^2}{16\pi^2} \right) [-2C_F + 2(1-\xi)C_F]$$

$$(2.136) \quad \delta_m = \frac{1}{\epsilon} \left(\frac{g^2}{16\pi^2} \right) [-6C_F]$$

$$(2.137) \quad \delta_3 = \frac{1}{\epsilon} \left(\frac{g^2}{16\pi^2} \right) \left[\frac{10}{3}C_A - \frac{8}{3}n_f T_F + (1-\xi)C_A \right]$$

$$(2.138) \quad \delta_{3c} = \frac{1}{\epsilon} \left(\frac{g^2}{16\pi^2} \right) \left[C_A + \frac{1}{2}(1-\xi)C_A \right]$$

$$(2.139) \quad \delta_{A^3} = \frac{1}{\epsilon} \left(\frac{g^2}{16\pi^2} \right) \left[\frac{4}{3}C_A - \frac{8}{3}n_f T_F + \frac{3}{2}(1-\xi)C_A \right]$$

$$(2.140) \quad \delta_{A^4} = \frac{1}{\epsilon} \left(\frac{g^2}{16\pi^2} \right) \left[-\frac{2}{3}C_A - \frac{8}{3}n_f T_F + 2(1-\xi)C_A \right]$$

$$(2.141) \quad \delta_{1c} = \frac{1}{\epsilon} \left(\frac{g^2}{16\pi^2} \right) [-C_A + (1-\xi)C_A]$$

5.3. Physical observables.

6. BRST symmetry

Recall that we have the following form of the path integral:

$$(2.142) \quad \mathcal{Z} = \int \mathcal{D}A_\mu^a \mathcal{D}\mathcal{B}^a \mathcal{D} \prod \delta(\partial^\mu A_\mu^a - \mathcal{B}^a)$$

$$(2.143) \quad \exp\left(i \int d^4x \operatorname{tr}(F_{\mu\nu}F^{\mu\nu} + \xi \operatorname{tr} C^2 + ibM(\mathcal{A})c)\right)$$

But recall this is no longer gauge invariant. It is however invariant under the BRST^{2.10} local fermionic symmetry. Note that the auxiliary field \mathcal{B} can be integrated out since it is only present in a quadratic exponential, and a delta function. So this is often simplified to:

$$(2.144) \quad \mathcal{Z} = \int \mathcal{D}A_\mu^a \exp\left(i \int d^4x \operatorname{tr}(F_{\mu\nu}F^{\mu\nu}) + 1/\xi (\partial^\mu A_\mu)^2 + ibM(\mathcal{A})c\right)$$

As before, this symmetry is given by some charge Q ^{2.11} except now Q is a fermionic scalar. In particular Q satisfies the anti-commutator relation: $\{Q, Q\} = 0$. In other words, Q_{brst} satisfies the super charge condition $Q^2 = 0$. So what were people doing to find this?

(1) Consider QED.

(a) First observe that if $\square\alpha = 0$, then there is a residual symmetry:

$$(2.145) \quad A_\mu \rightarrow A_\mu + \partial_\mu \alpha$$

(b) Observe that the equations of motion for c are $\square x = 0$, and $\square \bar{c} = 0$.

(c) If I define $A_\mu \rightarrow A_\mu + \theta \partial_\mu c$ where θ is an infinitesimal parameter of this transformation it is also a real Grassmannian which cannot depend on spacetime. This means $\theta^2 = 0$. Then we define this to be the action of Q :

$$(2.146) \quad (\theta Q) A_\mu$$

Note now that we have

$$(2.147) \quad \delta(\partial^\mu A_\mu)^2 = 2(\partial^\mu A_\mu)(\theta \square c)$$

and now we can define the variations of the ghost/antighost to compensate for this:

$$(2.148) \quad \delta \bar{c} = -\frac{\theta}{\xi} \partial^\mu A_\mu$$

so Q is indeed a symmetry. Note we didn't have to introduce any complicated variation of the ghost. Implicitly we set this to be 0. In summary, the Abelian action of Q is given by:

$$(2.149) \quad \delta A_\mu = \theta \partial_\mu c \quad \delta \bar{c} = -\frac{\theta}{\xi} \partial^\mu A_\mu \quad \delta c = 0$$

(2) For non-Abelian Yang-Mills, it turns out this same formalism basically works. We can guess

$$(2.150) \quad \delta A_\mu^a = \frac{\theta}{g_{\text{ym}}} D_\mu c^a \quad \delta \bar{c} = -\frac{1}{g} \frac{\theta}{\xi} \partial^\mu A_\mu^a$$

^{2.10} This stands for Becchi, Rouet, and Stora and Tyutin.

^{2.11} This was originally called s and is sometimes called Δ .

but it turns out this is not a good symmetry, because this object $M(\mathcal{A})$ is effectively $\partial^\mu D_\mu$ which clearly depends on \mathcal{A} . But now this symmetry will spit out some term like $\theta \bar{c} c c$ so we need to get rid of this cubic term. To fix this, postulate

$$(2.151) \quad \delta c^a = -\frac{1}{2} \theta f^{abc} c^b c^c$$

Note that if we didn't integrate out the auxiliary field \mathcal{B} , then $\delta \bar{c}^a = \theta \mathcal{B}^a$ and $\delta \mathcal{B}^a = 0$. This is the most elementary transformations where $Q^2 = 0$.

Generic gauge theories

Yang Mills is an example of a gauge theory, but this is really a broad family of theories. Even gravity can be viewed as one. Recall we found this so called BRST symmetry almost by accident in Yang Mills. We have this BRST charge Q where $Q^2 = 0$.

This is reminiscent of the differential geometric setup where we have some differential which squares to 0. So we have some manifold (M, ω) . Then we write the exterior derivative as d . Now we can ask what the Hilbert space of physical states is?

The best definition will be the following:

- (1) Construct \mathcal{H} for A_μ , and the ghost and anti-ghost c, b . This is a huge Fock space of all the fields. Here we even have the possibility that the ghosts can carry momentum.
- (2) Define an equivalence of states such that they are considered the same iff they differ by the action of Q on something:

$$(3.1) \quad |\varphi\rangle \cong |\varphi\rangle + Q|\chi\rangle$$

Then we claim that the cohomology of Q is the space of physical states.

1. Instantons

Much of this section is from [7]. See this for a more detailed treatment.

DEFINITION 3.1. An instanton is a classical solution to the equations of motion (up to the total derivative in S) with finite action.

Basically everything in this section works with the semi-classical approximation. So using Hamiltonian methods in quantum mechanics, we use WKB.

Say we have some φ^4 theory. So we have some Lagrangian:

$$(3.2) \quad \mathcal{L} = \frac{1}{2} \partial_\mu \varphi \partial^\mu \varphi - \frac{1}{2} m^2 \varphi^2 - \frac{g^2}{4!} \varphi^4$$

Classically, we have that $\varphi' = g\varphi$, so we have the classical Lagrangian:

$$(3.3) \quad \mathcal{L} = \frac{1}{g^2} \left(\frac{1}{2} \partial_\mu \varphi \partial^\mu \varphi - \frac{1}{2} m^2 \varphi^2 - \frac{1}{4!} \varphi^4 \right)$$

Now recall that we have another parameter to worry about. Namely, \hbar . Then there is a sense in which \hbar and g conspire to get $\mathcal{L}/\hbar = 1/(g^2 \hbar) (\dots)$, so when we take the $\hbar \rightarrow 0$ limit to get a semi-classical approximation, this is like taking a small g limit as well. We really just have one parameter: $g^2 \hbar$.

We might wonder at this point whether we get a complete picture of things with perturbation theory alone. This is in fact not the case. We will see this even in the basic theory of quantum mechanics for a single particle.

EXAMPLE 3.1. We have a simple example:

$$(3.4) \quad L = \frac{1}{2}x'^2 - V(x; g)$$

where $V(x; g) = 1/g^2 F(gx)$. and F is some Taylor expansion which begins with terms of order x^2 . It is a basic exercise in quantum mechanics to calculate the transmission amplitude through a potential barrier:

$$(3.5) \quad |T(E)| = \exp \left\{ -\frac{1}{\hbar} \int_{x_1}^{x_2} dx [2(V-E)]^{1/2} \right\} (1 + \mathcal{O}(\hbar))$$

We see, however, that transmission is not seen in any order of perturbation theory, because our expression vanishes more rapidly than any power of \hbar and therefore more rapidly than any power of g . So we need other methods to see these instanton solutions.

1.1. Quantum mechanics of the harmonic oscillator. We will be focusing on the path integral formalism. Recall that the path integral for euclidean time is written:

$$(3.6) \quad \langle x_f | e^{-HT/\hbar} | x_i \rangle = \mathcal{N} \int [dx] e^{-S_E/\hbar}$$

Now recall that the harmonic oscillator has $H|n\rangle = E_n|n\rangle$ where

$$(3.7) \quad H = \frac{p^2}{2} + V(x)$$

and we can write:

$$(3.8) \quad \langle c_f | e^{-HT/\hbar} | x_i \rangle = \sum_n e^{-E_n T/\hbar} \langle x_f | n \rangle \langle n | x_i \rangle$$

where

$$(3.9) \quad S_E = \int_{-T/2}^{T/2} dt \left[\frac{1}{2} \left(\frac{dx}{dt} \right)^2 + V(x) \right]$$

Now if we take

$$(3.10) \quad x(t) = \bar{x}(t) + \sum_n x_n x_n(t) \quad [dx] = \prod_n (2\pi\hbar)^{-1/2} dc_n$$

Now the euclidean equation of motion is given by:

$$(3.11) \quad \frac{\delta S_E}{\delta \bar{x}} = -\frac{d^2 \bar{x}}{dt^2} + V;(\bar{x}) = 0$$

then for x_n , we have

$$(3.12) \quad -\frac{d^2 x_n}{dt^2} + V''(\bar{x}) x_n = \lambda_n x_n$$

Now we can calculate

$$(3.13) \quad \langle x_f | e^{-HT/\hbar} | x_i \rangle = \mathcal{N} e^{-S(\bar{x})/\hbar} \prod_n \lambda_n^{-1/2}$$

So this is exact for the actual harmonic oscillator, but if we are using a generic potential there will be a factor of $(1 + \mathcal{O}(\hbar))$ tagged on.

Now the limit we are really interested in is $T \rightarrow \infty$. Now we know $\omega^2(\bar{x}) = V''(\bar{x})$ with boundary conditions $x_f = 0$ at $T/2$ and $x_i = 0$ at $-T/2$. This means $\omega^2 = V''(0)$, so $\mathcal{N} \det(\dots) = \left(\frac{\omega}{\pi\hbar}\right) e^{-\omega T/2}$ so $E = \hbar\omega/2$.

1.2. Quantum mechanics of a double well. In regular (real) time, we have two minima. We have two maxima in euclidean (imaginary) time. Now we can always just use Hamiltonian methods to treat this, but we will use the path integral formalism, so that we can more easily generalize to treat gauge theories. Assign the maxima in euclidean time to be at $\pm a$. Now we are interested in the two expectation values:

$$(3.14) \quad \left\langle \pm a \left| e^{-HT/\hbar} \right| \pm a \right\rangle \quad \left\langle \pm a \left| e^{-HT/\hbar} \right| \mp a \right\rangle$$

Then our goal is to find solutions to the classical euclidean equations of motion. Now looking at the equations of motion, we can see the following is true. For $E = 0$,

$$(3.15) \quad \frac{dx}{dt} = (2V)^{1/2}$$

Now we can solve this rather easily to get $t = t_1 + \int_0^x dx' (2C(x'))^{-1/2}$. So this is a pseudo particle. It seems to be a solution centered at some point in time. This is where the name instanton came from. We call the corresponding solution for $a \rightarrow -a$ an anti-instanton. The action here can be written:

$$(3.16) \quad S_E = \int dT \left[\frac{1}{2} \left(\frac{dx}{dt} \right)^2 + V \right] = \int dT \left(\frac{dx}{dt} \right)^2 = \int_{-a}^a dx (2V(x))^{1/2}$$

so in the large T approximation, $x \rightarrow a$, and $dx/dt = \omega(a-x)$.

Now from the differential equation, we can solve for $a-x$. In particular,

$$(3.17) \quad a-x \propto e^{-\omega t}$$

This is some decaying function, so $1/\omega$ is the size of the instanton.

We now make a few remarks:

- (1) For n widely separated instantons ($\Delta d \gg 1/\omega$) then $S = nS_E$ where n is the number of instantons and anti-instantons.
- (2) We already know the usual form of the determinant, and here we just have a correction K :

$$(3.18) \quad \left(\frac{\omega}{\pi\hbar} \right)^{1/2} e^{-\omega T/2} K^n$$

- (3) We just integrate over all centers.
- (4) Because of the double well structure, every instanton has to be following by an anti-instanton. This means if we are doing calculations, we will have corrections:

$$(3.19) \quad \left\langle -a \left| e^{-HT/\hbar} \right| -a \right\rangle = \left(\frac{\omega}{\pi\hbar} \right)^{1/2} e^{-\omega T/2} \sum_{\text{nodd}} \frac{(Ke^{-S_E/\hbar T})^n}{n!} [1 + \mathcal{O}(\hbar)]$$

We have the expression:

$$(3.20) \quad \left\langle \pm \left| e^{-HT/\hbar} \right| -a \right\rangle = \frac{1}{2} \left(\frac{\omega}{\pi\hbar} \right)^{1/2} e^{-\omega T/2} \left(e^{K \exp(-S_E/\hbar)T} \mp e^{-K \exp(-S_E/\hbar)T} \right)$$

Now the important point is, that the energy itself is modified by barrier penetration.

$$(3.21) \quad |\langle + | \pm \rangle|^3 = \frac{1}{2} \left(\frac{\omega}{\pi\hbar} \right)^{1/2} \rightarrow E_{\pm} = \frac{1}{2} \hbar\omega \pm \hbar K e^{-S_E/\hbar}$$

So the final term is non-perturbative.

1.3. Quantum mechanics of a sinusoidal potential. Now consider a sinusoidal potential $e^{ix_j\theta}$ where x_j is the j th minimum of the real time potential. We obtain

$$(3.22) \quad \langle j_+ | e^{-HT/\hbar} | j_- \rangle = \sum_{n=0}^{\infty} \sum_{\bar{n}=0}^{\infty} \frac{!}{n!\bar{n}!} \left(K e^{-S_E/\hbar T} \right)^{n+\bar{n}} \delta_{n-\bar{n}-j_++j_-} \left(\frac{\omega}{\pi\hbar} \right)^{-1/2} e^{-\omega T/2}$$

where n is the number of instantons and \bar{n} is the number of anti-instantons. If we use the identity

$$(3.23) \quad \delta_{ab} = \int_0^{2\pi} d\theta e^{i\theta(a-b)}/2\pi$$

the sum becomes two independent exponential series, and we can write:

$$(3.24) \quad \langle j_+ | e^{-HT/\hbar} | j_- \rangle = \sqrt{\frac{\omega}{\pi\hbar}} e^{-\omega T/2} \int_0^{2\pi} e^{i(j_- - j_+)\theta} \frac{d\theta}{2\pi} \exp\left(2KT \cos \theta e^{-S_0/\hbar}\right)$$

As such, the continuum of energy eigenstates is labeled by θ . The energy eigenvalues themselves are given by:

$$(3.25) \quad E(\theta) = \frac{\hbar\omega}{2} + 2\hbar K \cos \theta e^{-S_0/\hbar}$$

and

$$(3.26) \quad \langle \theta | j \rangle = \sqrt{\frac{\omega}{\pi\hbar}} (2\pi)^{-1/2} e^{ij\theta}$$

1.4. Instantons in SU(2) gauge theory in 4D. Now the real object of study here is the finite contribution to the action. For us to have finite action, there are some constraints that show up. We can refactor $g \equiv g_{\text{ym}}$ out front:

$$(3.27) \quad \mathcal{L} = \frac{1}{g^2} \text{Tr} (F_{\mu\nu}^a F_{\mu\nu}^b)$$

now F needs to fall off like $F \sim \mathcal{O}(r^{-3})$ and A needs to fall off like $A \sim g\partial_\mu g^{-1} + \mathcal{O}(r^{-2})$ from the derivative, and since A is only defined up to a gauge transformation. So the problem with the first term of this asymptotic expansion for A , is that we cannot gauge away globally. This is because, if we try to perform a gauge transformation to gauge these away, these fall into different classes. But not all of these can be continuously deformed such that they all go to the identity.

In any case, we have a boundary space, which maps into the Gauge group. so we are looking for non-trivial continuous mappings.

EXAMPLE 3.2. If our gauge group is $U(1)$, and we are in two dimensions, the one point compactification of the boundary is S^1 . So are there any mappings where $S^1 \rightarrow S^1 \cong U(1)$? Yes, we can have $g^{(0)} = 1$. But globally we can't make this equal 1, $g^{(I)} = e^{i\theta}$. Similarly,

$$(3.28) \quad g^{(n)} = \left(g^{(I)}\right)^n = e^{in\theta}$$

for $n \in \mathbb{Z}$.

EXAMPLE 3.3. In four dimensions, our boundary looks like S^3 , and we can map $S^3 \rightarrow S^1$. So we don't have topological obstructions for a $U(1)$ gauge field in 4 dimensions.

So we don't have to worry for Abelian theories in 4 dimensions. But we do have to worry for non-abelian theories.

EXAMPLE 3.4. Now consider $SU(2)$ in 4 dimensions. We have

$$(3.29) \quad g = \nu \Pi + i \vec{b} \vec{\sigma}$$

where σ denotes the pauli matrices, and ∂ depends on spacetime. Indeed we have the same as $U(1)$ in 2 dimensions:

$$(3.30) \quad g^{(0)} = 1 \quad g^{(I)}(x) = \frac{x_0 + i \vec{x} \vec{\sigma}}{r} \quad g^{(\nu)}(x) = \left(g^{(I)}\right)^\nu$$

For $U(1)$ in two dimensions,

$$(3.31) \quad \nu = \frac{1}{2\pi} \int_0^{2\pi} d\theta d \frac{dg^{-1}}{d\theta} = \frac{1}{2\pi} \int d^2x \mathcal{E}_{\mu\nu} F_{\mu\nu}$$

For $SU(2)$, where θ_i parameterizes S^3 ,

$$(3.32) \quad \nu = -\frac{1}{24\pi^2} \int d\theta_1 d\theta_2 d\theta_3 \text{Tr}(\epsilon^{ijh} d\partial_i g^{-1} \times g \partial_j g^{-1} g \partial_\mu g^{-1})$$

$$(3.33) \quad = \frac{1}{32\pi^2} \int \partial^4 x \tilde{F} F$$

$$(3.34) \quad \tilde{F} = \frac{1}{2} \epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma}$$

Now recall that if we are in the axial gauge, we have no ghosts. So in this setting, we can explore instantons without concerning ourselves over ghosts. Of course we can always go to the Faddeev-Popov setting and do some calculations in a simpler fashion, the axial Gauge will suit our purpose well. So here $A_3 = 0$. Then we have

$$(3.35) \quad F(V, T, n) = \mathcal{N} \int [dA] e^{-S} \delta_{\nu n}$$

Now we find, when we impose different instanton solutions:

$$(3.36) \quad F(V, T_1 + T_2, n) = \sum_{n_1+n_2=n} F(V, T_1, n_1) \otimes F(V, T_2, n_2)$$

Now we can basically perform a Fourier transform to get:

$$(3.37) \quad F(V, T, \theta) = \sum_n e^{in\theta} F(V, T, n) = \mathcal{N} \int [dA] e^{-S} e^{i\nu\theta}$$

which means we can factor:

$$(3.38) \quad F(V, T_1 + T_2, \theta) = F(V, T_1, \theta) F(V, T_2, \theta)$$

and we know

$$(3.39) \quad \nu = \frac{\theta}{32\pi^2} \int F \tilde{F}$$

2. BRST quantization of gauge theories

A good additional reference for this section is chapter 4.3 of [20]. We will first treat a general gauge theory. We will have fields φ_i where i is a single index standing for μ, a , and x . In particular, for a YM theory this will be $A_\mu^a(x)$. We will call the gauge transformations δ_α . Here α collectively refers to a and x . For YM, this will of course be $\delta A_\mu^a(x) = D_\mu \beta^a(x)$. We will assume that the infinitesimal gauge transformations form a closed Lie algebra. This means that we essentially have two assumptions: $[\delta_\alpha, \delta_\beta] = f^\gamma_{\alpha\beta} \delta_\gamma$, and that $f^\gamma_{\alpha\beta}$ are constants. These two assumptions can however fail. This can even happen in gravity or super-gravity.^{3.1} Now we choose a gauge fixing condition $F^A(\varphi_i) = 0$. For example, we have seen the Lorentz gauge fixing condition to be $0 = \mathcal{F} = \partial^\mu A_\mu^a$ so this collective index A stands for both μ and a in this example.

Now we have three objects, the φ_i , δ_α , and $F^A(\varphi_i)$. This admits a path integral representation:

$$(3.40) \quad \int \frac{\mathcal{D}\varphi_i}{\text{Vol}(\mathcal{G})} e^{-S_1(\varphi_i)}$$

which is somewhat ill defined, since we are taking the volume of the infinite group \mathcal{G} to cancel the redundancies in the description. Recall in the case of YM, $S_1 = 1/(2g^2) \int \text{tr}(FF)$. Now we claim that we are defining, and disentangling this expression to get:

$$(3.41) \quad \int \mathcal{D}\varphi_i \mathcal{D}B_A \mathcal{D}b_A \mathcal{D}c^\alpha e^{-S_1 - S_2 - S_3 - S_4}$$

where c will be the so-called ghost, b the anti-ghost if we are assuming the system is fermionic, then these will be fermions. B is the auxiliary field boson. So in addition to the three objects from before, we have now introduced c^α, b_A, B_A . Two of the additional terms in the action are:

$$(3.42) \quad S_2 = -iB_A F^A(\varphi) \quad S_3 = b_A \delta_\alpha F^A(\varphi) c_A$$

We will call the BRST charge $Q_{\text{BRST}} \equiv Q$. This generated a global symmetry. The action is given by:

$$(3.43) \quad \delta_B \varphi_i = -i\epsilon c^\alpha \delta_\alpha \varphi_i \quad \delta_B c^\alpha = \frac{i}{2} \epsilon f^\alpha_{\beta\gamma} c^\beta c^\gamma$$

$$(3.44) \quad \delta_B B_A = 0 \quad \delta_B b_A = B_A$$

where ϵ is not in any way spacetime dependent. This is just a single Grassmannian parameter, so $\epsilon^2 = 0$. The final term in the action is then given by:

$$(3.45) \quad \epsilon S_4 = \delta_B (b_A B_B M^{AB})$$

where M is an arbitrary matrix of constants independent of any fields. These terms seem arbitrary at this point, but after we discuss the space which we will identify with the physical states of the system, we will understand why these terms make sense.

In any case, we now have four fields, which split into two subcollections φ_i with c^α and b_A with B_A . This is because b_A and B_A comprise the smallest collection of fields on which $Q^2 = 0$ without any equations of motion. In addition to Q , we also have a $U(1)$ symmetry, given by the ghost number. In particular, we denote this by U and take

$$(3.46) \quad U(\varphi_i) = 0 \quad U(c^\alpha) = 1 \quad U(b_A) = -1 \quad U(B) = 0$$

This is a global symmetry with an anomaly.

^{3.1} We can treat this with the BRST-BV formalism.

First define

$$(3.47) \quad \psi := b_A Q F^A(\varphi)$$

where Ψ is often called the gauge fixing fermion. This gives us

$$(3.48) \quad i\epsilon(S_2 + S_3)$$

In other words, the $S_2 + S_3$ part of the action is the BRST transformation of something. In fact, we have this for S_4 as well. Recall that

$$(3.49) \quad \epsilon S_4 = \delta_B(b_A B_B M^{AB})$$

where M is an arbitrary matrix of constants independent of any fields. Clearly the constants are BRST invariant, and B_B is invariant, so the action of Q gives us a quadratic term $B_A B_B M^{AB}$. Now notice we have added something which is Q of something, so it doesn't change the physics. S_1 wasn't like this, it wasn't Q of anything, but this is okay because it was a starting point for us.

Why is $Q^2 = 0$ physically meaningful? First we have the question of what the Hilbert space of physical states is. We know if we just take the Fock space of φ_i , this is too large. What about the Fock space \mathcal{H} of φ_i, b, c ? Well \mathcal{H} is clearly even larger, so we need to identify conditions for $|\psi\rangle \in \mathcal{H}$ to be physical. In particular, we know physics doesn't change when we change $F^A \rightarrow F^A + \delta_G F^A$. This means

$$(3.50) \quad 0 = \epsilon \delta_G \langle f|i\rangle = i \langle f|\delta_B(b_A \delta F^A)|i\rangle = -\epsilon \langle F|\{Q, b_A \delta_G F^A\}|i\rangle$$

so $Q|i\rangle = 0$ and $Q|f\rangle = 0$. This is clearly necessary, but this is in fact not sufficient. To see this, consider any state $\langle f|$ such that $\langle F|Q = 0$, and choose $|i\rangle$ such that there is some $Q|\chi\rangle$, meaning $Q|i\rangle = 0$. Now we have:

$$(3.51) \quad \langle f|i\rangle = \langle f|Q|\chi\rangle = 0$$

Therefore we need to identify

$$(3.52) \quad |\psi\rangle \cong |\psi\rangle + Q|\chi\rangle$$

for any $\chi \in \mathcal{H}$. More technically $\mathcal{H}_{\text{physical}}$ is the Q -cohomology on \mathcal{H} . States that are annihilated by Q are called Q -closed, and then $\mathcal{H}_{\text{physical}}$ can be said to be the Q -closed states modulo Q -exact states.

2.1. Cohomology. An abstract operator d which squares to 0 gives a more general cohomology theory. Start with some differentiable manifold M of dimension d , with some coordinates x^i . Take a special subclass of all possible tensors. In particular take covariant, rank- p , fully anti-symmetric tensors for $p \in \{0, \dots, d\}$. These are called closed forms.

$$(3.53) \quad \omega^{(p)} = \omega_{i_1, \dots, i_p}(x) dx^{i_1} \wedge \dots \wedge dx^{i_p}$$

We also have a differential operator $d: \omega^{(p)} = (d\omega)^{(p+1)}$ where $d^2 = 0$. Now if we define

$$(3.54) \quad V_{\text{closed}} = \{v | dv = 0\} \quad V_{\text{exact}} = \{v | \exists w \text{ s.t. } v = vw\}$$

We can define the cohomology of d to be

$$(3.55) \quad H(V, d) = H(V) = V_{\text{closed}}/V_{\text{exact}}$$

We can then refine this to be restricted to forms of degree p , and then this would refine our definition to the so called p th cohomology group $H^p(V)$. For our manifold M , this is called the de Rham cohomology $H^p(M)$.

EXAMPLE 3.5. We can now calculate the de Rham cohomology of the two-torus. In particular, $H^0(T^2) = \mathbb{R}$ and $H^2(T^2) = \mathbb{R}$. Any constant is a generator of H^0 , and we can check there are no others. We have that $d(dx \wedge dy) = 0$ so this is a non-trivial generator of H^2 , and we can check there are no others. We also have that $H^1(T^2) = \mathbb{R}^2$ where the basis is dx and dy . These are d of something locally, but not globally.

2.2. BRST cohomology. Returning to our BRST charge Q , we have the natural treatment of Q as the differential, and \mathcal{H} as V . The degree p of the cohomology is analogous to the ghost number of our theory. This is self consistent with the fact that d raises the degree of the form by 1, and Q raises the ghost number by 1. Note that as in the general case, we can label our cohomology groups by an integer. In this case this corresponds to the ghost number.

Now that we see how important this condition is, we justify why Q should square to 0. In particular, we see that don't have to actually assume this artificially. We do assume that Q is conserved under Hamiltonian evolution, and between gauge fixing conditions:

$$(3.56) \quad 0 = [Q, \delta H] = [Q, \{Q, b_A \delta_G F^A\}] = [Q^2, b_A \delta F^A]$$

so Q^2 commutes with any such object. This doesn't mean $Q^2 = 0$, but it does mean it is a constant. But now we realize that there is no constant bosonic object with ghost number 2, so it better be 0.

EXAMPLE 3.6. Consider the system given by a point like particle where we have the fields $\varphi^i \hat{=} X^\mu(\tau)$ and $e(\tau)$. The metric here is given by

$$(3.57) \quad g_{\alpha\beta}(\sigma) = e_\alpha^A(\sigma) e_\beta^B(\sigma) \eta_{AB}$$

Now the most generic gauge invariant action we can cook up is:

$$(3.58) \quad S = \int d\tau \left(\frac{1}{2} e X'^\mu X'^\nu \eta_{\mu\nu} + \frac{1}{2} e m^2 \right)$$

note $\sqrt{|g|} = e$. Then the δ_α are some single dimensional differentials. In particular,

$$(3.59) \quad \delta X^\mu(\tau) = i\epsilon c(\tau) X'^\mu(\tau)$$

following the strategy of BRST. Similarly,

$$(3.60) \quad \delta e = i\epsilon(c'e)$$

Now to get that $Q^2 = 0$, we need to insist on the sense in which the ghost transforms:

$$(3.61) \quad \delta c = i\epsilon c c'$$

Now to write down a decent kinetic term for the ghost, we need to introduce an auxiliary field B and antighost b . In particular these transform as:

$$(3.62) \quad \delta B(\tau) = 0 \quad \delta b(\tau) = B(\tau)$$

Now recall in the generic BRST treatment we had $S_{\text{BRST}} = S_1 + S_2 + S_3 = S_1 + \{Q, \Psi\}$
Now if we take the sensible gauge fixing condition $e = 1$, then

$$(3.63) \quad S_2 = \int iB(e-1) \quad S_3 = - \int eb'c$$

Now integrating out the auxiliary field, we get $e = 1$ and no B in the path integral. So

$$(3.64) \quad S = \int d\tau \left(\frac{1}{2} (X'^\mu)^2 + \frac{1}{2} m^2 - b'c \right)$$

Not this is a good Gaussian, we can just solve it. We consider the physical Hilbert space. Take $Q = cH$ where $H = 1/2 (p^2 + m^2)$. We can write:

$$(3.65) \quad [p^\mu, X^\nu] = -i\eta^{\mu\nu} \quad \{b, c\} = 1$$

which shows us that

$$(3.66) \quad \mathcal{H} = \mathcal{H}_{p,x} \otimes \mathcal{H}_{b,c}$$

where $\mathcal{H}_{p,x}$ is given by the kets $|k^\nu\rangle$ such that

$$(3.67) \quad p^\mu |k^\nu\rangle = k^\mu |k^\nu\rangle$$

So it consists of states labeled by a certain momentum. The subtlety comes from the interpretation of the $\mathcal{H}_{b,c}$. In particular, this is a two dimensional Hilbert space given by

$$(3.68) \quad \mathcal{H}_{b,c} = \text{Span} \{|\downarrow\rangle, |\uparrow\rangle\}$$

where we declare:

$$(3.69) \quad b|\downarrow\rangle = 0 \quad b|\uparrow\rangle = |\downarrow\rangle$$

$$(3.70) \quad c|\downarrow\rangle = |\uparrow\rangle \quad c|\uparrow\rangle = 0$$

Now we have:

$$(3.71) \quad Q|k, \downarrow\rangle = \frac{1}{2} (k^2 + m^2) |k, \uparrow\rangle \quad Q|k, \uparrow\rangle = 0$$

Now following the preceding treatment, we declare the Q -closed states to be all $|k, \downarrow\rangle$ with $k^2 = -m^2$ and all $|k, \uparrow\rangle$. Then take none of the $|k, \downarrow\rangle$ to be Q -exact, and take $|k, \uparrow\rangle = Q|\chi\rangle$ iff $k^2 + m^2 \neq 0$. Now we can simple take

$$(3.72) \quad \mathcal{H}_{\text{physical}} := \mathcal{H}_{\text{closed}} / \mathcal{H}_{\text{exact}}$$

which consists of $|k, \downarrow\rangle$ such that $k^2 + m^2 = 0$ and $|k, \uparrow\rangle$ such that $k^2 + m^2 \neq 0$. Then we have ghost numbers $\mathcal{U}(|\downarrow\rangle) = -1/2$ and $\mathcal{U}(|\uparrow\rangle) = 1/2$ so we can see that

$$(3.73) \quad \mathcal{H}_{\text{physical}} = \mathcal{H}_{\text{physical}}^{-1/2} \oplus \mathcal{H}_{\text{physical}}^{1/2}$$

Now we add the condition that $b|\text{physical}\rangle = 0$. So we see that the rules usually basically work, but we sometime shave to add this additional condition to get rid of some degeneracies. This is referred to as equivariant cohomology.

3. Topological quantum field theories

There is some ambiguity as to what topological means in this context. We will mostly focus on cohomological quantum field theories. In particular, we will concern ourselves mostly with topological YM in four dimensions. Other examples of cohomological quantum field theories are given by the so-called topological σ -models in $1+1$ dimensions. The fields and transformations for this theory are given by:

$$(3.74) \quad \varphi : X^i(\tau, \sigma) \quad \delta_\alpha : \delta X^i(\tau, \sigma) = \xi^i(\tau, \sigma)$$

The action is given by:

$$(3.75) \quad S_1 = \int \omega_2 = T + \{W, \Psi\}$$

where T is topologically invariant.

Another example of a cohomological QFT is given by topological quantum mechanics. This is a theory in $0 + 1$ dimensions. The fields are given by

$$(3.76) \quad X^i(t) \quad \delta : \delta X^i = \xi^i(t)$$

Then performing BRST gives us that this is automatically $N = 2$ super-symmetric quantum mechanics. For more on this theory, see [1].

There are also theories considered to be topological which are not cohomological. In particular we will discuss the s-called Chern-Simons theories with action given by:

$$(3.77) \quad S = \int \text{tr} \left(A \wedge dA + \frac{2}{3} A \wedge A \wedge A \right)$$

4. Topological Yang-Mills gauge theories

Consider M^4 a 4 dimensional Riemannian manifold with a metric $g_{\mu\nu}(x^\mu)$ and take $G = \text{SU}(2)$. Then our fields are given by

$$(3.78) \quad A_\mu^a(x) \quad \delta A_\mu^a = f_\mu^a(x) + D_\mu f^a(x)$$

So we state that any deformation of A is a gauge symmetry. This means we have no propagating degrees of freedom, so no “gluons” in this theory. This is what we mean by topological here. We might wonder what are chances are to treat this sort of theory perturbatively and come up with Feynman rules? In fact, things get even worse when we realize that the best invariant action we can have is:

$$(3.79) \quad S_{\text{tym}} = \int_{M^4} \text{tr}(F \wedge F)$$

since $\text{tr}(F \wedge F)$ is a total derivative and topological invariant. Here, topological means that it is independent of local changes in $g_{\mu\nu}(x)$.

WARNING 3.1. We have met two different notions of a topological theory. First that the theory has no local degrees of freedom, and therefore no gluons. The second, is that the theory is independent of $g_{\mu\nu}(x)$. We will also see that there is a notion that a theory is topological if it can be written without $g_{\mu\nu}$.

EXAMPLE 3.7. Take $M = \mathbb{R}$, then we seek to show this action is a topological invariant. We have that

$$(3.80) \quad S = \int_{\mathbb{R}^4} \text{tr}(F \wedge F) = \int_{S^3} \text{tr} \left(A \wedge dA + \frac{2}{3} A \wedge A \wedge A \right)$$

where the three form in the second integral is called the *Chern-Simons form*. This form clearly doesn’t care about the metric on the 3-sphere. This action give us the Chern-Simons gauge theory. This is a special case for $2 + 1$ dimensions.

Now recall that an instanton is a solution of the equation of motion of YM in a sector with instanton number 1. Then a multi-instanton has instanton number n . We have

$$(3.81) \quad \mathcal{Z} = \int \frac{\mathcal{D}A}{\text{Vol}(\mathcal{G})} d^{-S(A)}$$

Now to take a $\hbar \ll 1$ semi-classical approximation, we want to find the saddle points. This yields

$$(3.82) \quad \mathcal{Z} \approx \sum_n e^{-\hbar S(A_0)} (\det + \mathcal{O}(\hbar))$$

where n runs for instanton numbers. Now instantons minimize the action, so we get

$$(3.83) \quad 0 \leq \int d^4x \operatorname{tr} \left(F_{\mu\nu} \pm \frac{1}{2} \epsilon_{\mu\nu\sigma\rho} F^{\sigma\rho} \right)^2 = \int d^4x \operatorname{tr} \left(F_{\mu\nu}^{(\pm)} \right)^2$$

so

$$(3.84) \quad S(A) \sim \int d^4x + (F_{\mu\nu} F^{\mu\nu}) \geq \left| \int F \wedge F \right| \sim |n|$$

so the minima correspond to satisfying the self-dual and anti-self-dual conditions $F_{\mu\nu}^{(\pm)} = 0$. We call this the instanton equation.

In any case, we set

$$(3.85) \quad S_1 = \int \operatorname{tr} \{ F \wedge F \}$$

Our plan will be to consider a trivial Gauge bundle, and set $S_{\text{tym}} = 0$. Now use the BRST machinery to add further terms to S , and see if we can get to a point where we can perform perturbative calculations.

So we have chosen our fields and our gauge transformations, now we need to determine our gauge fixing conditions, and then we can determine the ghosts and auxiliary field. So we need some insight in order to understand how we can insist on a useful gauge fixing condition. To find this, we can take advantage of our knowledge of instantons. So what is F^A ? Well if we have two choices, such that they can be interpolated between using Feynman rules, then we know these theories are the same. But if there is some choice where we cannot interpolate arbitrarily, then these two “disconnected” conditions are not physically equivalent. Recall that we can decompose $F_{\mu\nu} = F_{\mu\nu}^+ + F_{\mu\nu}^-$ into the self dual, and anti-self dual pieces. Recall that instantons are given by either $F^- = 0$ or $F^+ = 0$. This is called the instanton equation in physical YM. Ed Witten was smart to propose the gauge fixing condition: $F_{\mu\nu}^+ = 0$.

So now as we have seen,

$$(3.86) \quad S_{\text{tym}} = S_1 + \{Q, \Psi\}$$

But now we notice that we have to add ghosts, anti-ghosts, and their ghosts and anti-ghosts, and things can get quite complicated. Embed this into a class of specific types of topological theories.

WARNING 3.2. We have seen three notions of topological theories. There is another class of theories, known as cohomological topological quantum field theories. They always start with a topologically invariant action S_1 , which gives no equation of motion, and are then treated with the BRST prescription. This is detailed in the paper [27].

We call the primary ghost ψ_μ^a :

$$(3.87) \quad [Q, A_\mu^a] = \psi_\mu^a$$

which will not be covariant under the BRST charge Q . So we have

$$(3.88) \quad \{Q, \psi_\mu^a\} = -D_\mu \varphi^a \quad [Q, \varphi^a] = 0$$

where φ^a is bosonic and $\mathcal{U}(\varphi^a) = 2$. Note that the way we have formulated things, Q^2 squares to a YM gauge transformation, and not zero. But we can just introduce a ghost for this symmetry as well, and get the augmentations:

$$(3.89) \quad [Q, A_\mu^a] = \psi_\mu^a + D_\mu c^a \quad \{Q, \psi_\mu^a\} = -D_\mu \varphi^a + [c, \psi_\mu]^a \quad [Q, \varphi^a] = 0 + [x, \varphi]^a$$

This resolution gives us ghosts for ghosts. So for any cohomological quantum field theory, we have that $[Q, \text{FIELD}] = \text{GHOST}$ and $\{Q, \text{GH}\} = \delta_\varphi^{\text{GAUGE}}(\text{FIELD})$ and $[Q, \varphi] = 0$ then Q^2 is a gauge transformation.

This gets quite complicated, and we haven't even treated anti-ghosts yet. We have two choices to treat this. If we have the full treatment, clearly $Q^2 = 0$, but we can simply by leaving out the terms concerning c^a , then the theory still has YM symmetry implicitly, and we just keep track of the fact that Q^2 is some YM gauge transformation.

Now we define the observable:

$$(3.90) \quad \mathcal{O}_{k,0}(x) = \text{Tr}(\varphi^k(x))$$

In particular, we have ghost number $\mathcal{U}(\mathcal{O}_{k,0}) = 2k$. The 0 tells us this is a 0-form.

The idea is, we can take

$$(3.91) \quad d\mathcal{O}_{k,0} = \{Q, \mathcal{O}_{k,1}(x)\}$$

where

$$(3.92) \quad \mathcal{O}_{k,1} = \kappa \text{Tr}(\varphi^{k=1} \psi_\mu) dx^\mu$$

Now we can continue doing this until we exhaust the dimension of spacetime. This gives us:

$$(3.93) \quad d\mathcal{O}_{k,4} = 0$$

So we have $\mathcal{O}_{k,i}(x)$ for $i \in \{0, 1, 2, 3, 4\}$. Now for some cycle C , we can integrate

$$(3.94) \quad \int_C \mathcal{O}_{k,1}(x)$$

which has ghost number $2k - 1$ and eventually get

$$(3.95) \quad \int_M \mathcal{O}_{k,4}$$

which is a candidate deformation of S . Now we can take

$$(3.96) \quad S = S_1 + \{Q, \Psi\} + \sum_k t_k \int \mathcal{O}_{k,4}$$

but to do this we need to specify auxiliary fields $B_{\mu\nu}^{(\pm)} \chi_{\mu\nu}^{(\pm)}$ and antighost for ghost. Actually we will have two sets: (λ, η) where λ has the quantum numbers of φ , and $\mathcal{U}(\lambda) = -2$. The second set will be (χ, H) where χ is a fermionic self-dual 2-tensor, and H is the auxiliary bosonic field. Finally, $\mathcal{U}(\chi) = -1$ and $\mathcal{U}(H) = 0$. We now postulate that the action is:

$$(3.97) \quad S = \frac{1}{e^2} \int \{Q, V\} d^4x$$

where we have dropped the static term S_1 to focus on the terms which give us our dynamics in the theory and

$$(3.98) \quad \Psi = \int V d^4x \quad V = \sqrt{g} \text{Tr} \left\{ -2\chi_{\alpha\beta} \left[H^{\alpha\beta} - \frac{1}{2} (F^{\alpha\beta} - F^{*\alpha\beta}) \right] - D_\alpha \lambda \psi^\alpha \right\}$$

where we recall that

$$(3.99) \quad F^A(\varphi) : F_{\mu\nu}^\pm = F_{\mu\nu} \pm F_{\mu\nu}^* \quad F_{\mu\nu}^* = \frac{1}{2} \epsilon_{\mu\nu\sigma\rho} F^{\sigma\rho}$$

Now we can write the leading terms

$$(3.100) \quad S = \frac{1}{e^2} \int \sqrt{g} \text{Tr} \left[\frac{1}{8} (F_{\alpha\beta} F_{\alpha\beta}^*)^2 + D_\alpha \lambda D^\alpha \varphi - D_\alpha \eta \psi^\alpha + \dots \right]$$

Now we can recognize the terms here in the first term with the square, as being of the form $F_{\alpha\beta}F^{\alpha\beta}$ which looks like the YM lagrangian, and then terms like $F_{\alpha\beta}F^{*\alpha\beta} = F \wedge F$ which are topological terms. The next terms show us that fermions have their own statistics, which is as it should be, since they are a result of this BRST treatment. There are of course many more terms, which are not of as much interest to us. Now we have the fundamental observation originally made by Witten, that this looks very much like $N = 2$ supersymmetric (SUSY) YM.

Now we have some properties of this action:

- (1) \mathbb{Z} and correlation functions for the observables are independent of $g_{\mu\nu}(x)$.

PROOF. If we change the metric $g_{\mu\nu} \rightarrow g_{\mu\nu} + \delta g_{\mu\nu}$ then we get $\Psi \rightarrow \Psi + \delta\Psi$ so $\delta S = \{Q, \delta\Psi\}$ but since this is BRST exact, this doesn't change the observables. \square

- (2) This is one loop exact. In other words, we can take $e \rightarrow d + \delta e$ so we can take the semi-classical limit and get the correct answer for the entire correlation function. This is closely related to the notion of localization in topological QFT. To compute this semi-classical limit we need the saddle points. These are in fact given by the instanton equations $F_{\mu\nu}^{\pm} = 0$. Now we can take

$$(3.101) \quad n = \frac{1}{64\pi^2} \int F \wedge F$$

Now from the Riemann-Roch theorem we get that the dimension of the moduli space of instantons $\dim \mathcal{M}_n$ is given by the number of fermion zero modes of $\mathcal{U} = 1$ minus the number of zero modes of $\mathcal{U} = -1$. Now on the space associated with $F_{\mu\nu}^+$ we get:

$$(3.102) \quad D_{\alpha}\delta A_{\beta} - D_{\beta}\delta A_{\alpha} - \epsilon_{\alpha\beta\gamma\delta}D^{\gamma}\delta A^{\delta} = 0$$

In addition, if we replace δA_{α} with ψ_{α} then we get the equation of motion from the action. So the number of independent solutions of ψ_{α} is a good starting prediction for the dimensionality of this space.

5. Super-symmetry

5.1. Features of $N = 1$ Yang-Mills. A lot of this section will be only considering superspace. The standard reference for this subject is [24]. Note that we will be using the notation

$$(3.103) \quad \sigma_{\alpha,\dot{\alpha}}^m \gamma_m = \gamma_{\alpha,\dot{\alpha}}$$

The starting point of this theory is the super-Poincaré algebra. Just like we would start our consideration of field theory with the Lorentz algebra, we start this consideration of super-symmetry considering super-Poincaré algebra. This is just the normal rotations boosts and translations with the added supersymmetry(SUSY) generators $Q_{\alpha}, \bar{Q}_{\dot{\alpha}}$. These are spinorial fermionic objects. We insist that:

$$(3.104) \quad \{Q_{\alpha}, \bar{Q}_{\dot{\alpha}}\} 2i\sigma_{\alpha,\dot{\alpha}}^m \partial_m$$

The actual realization of these is given by:

$$(3.105) \quad Q_{\alpha} = \frac{\partial}{\partial\theta^{\alpha}} - i\sigma_{\alpha,\dot{\alpha}}^m \bar{\theta}^{\dot{\alpha}} \partial_m$$

The representations of this algebra are superfields classified by superspin. Superfields are functions on a supermanifold $\mathbb{R}^{4,n}$ with coordinates:

$$(3.106) \quad z^A = \left(x^m, \theta^\alpha, \bar{\theta}^{\dot{\alpha}} \right)$$

then superfields will effectively be functions on these coordinates: $\Phi = \Phi(x, \theta, \bar{\theta})$. Typically we then take a ‘‘Taylor expansion’’ in the Fermionic Grassmann coordinates:

$$(3.107) \quad \Phi(x, \theta, \bar{\theta}) = \varphi(x) + \theta^\alpha \psi_\alpha(x) + \bar{\theta}_{\dot{\alpha}} \bar{\psi}^{\dot{\alpha}}(x) + \theta^2 F(x) + \dots + \theta^2 \bar{\theta}^2 D(x)$$

the coefficient functions are called the component fields. Then we define the action as follows:

$$(3.108) \quad \delta_\epsilon \Phi = (\epsilon Q + \bar{\epsilon} \bar{Q}) \Phi$$

We have the component projection:

$$(3.109) \quad \varphi(x) = \Phi(x, 0, 0) = \Phi|_{\theta=0} \quad Q_\alpha \Phi| = \frac{\partial}{\partial \theta^\alpha} \Phi| = \psi_\alpha(x)$$

$$(3.110) \quad \delta_\epsilon \Phi| = \delta_\epsilon \varphi = \left(\epsilon^\alpha Q_\alpha + \bar{\epsilon}_{\dot{\alpha}} \bar{Q}^{\dot{\alpha}} \right) \Phi| = \epsilon^\alpha \psi_\alpha + \bar{\epsilon}_{\dot{\alpha}} \bar{\psi}^{\dot{\alpha}}$$

Note however that this gives us:

$$(3.111) \quad \delta_\epsilon (Q_\alpha \Phi) \neq Q_\alpha \delta_\epsilon \Phi$$

To resolve this, we introduce the super covariant derivative:

$$(3.112) \quad \{D_\alpha, Q_\beta\} = \{D_\alpha, \bar{Q}_\beta\} = \dots = 0 \quad \{D_\alpha, \bar{D}_{\dot{\alpha}}\} = -2i\partial_{\alpha, \dot{\alpha}}$$

$$(3.113) \quad D_\alpha = \frac{\partial}{\partial \theta^\alpha} + i\bar{\theta}^{\dot{\alpha}} \partial_{\alpha \dot{\alpha}}$$

Now we use this to consistently define components of a superfield to be:

$$(3.114) \quad \varphi \equiv \Phi| \quad \psi_\alpha \equiv D_\alpha \Phi| \quad F \equiv D^2 \Phi| \quad \dots$$

We will also use this to covariantly constrain superfields. In particular, we can consider a real superfield $V = \bar{V}$:

$$(3.115) \quad V(x, \theta, \bar{\theta}) = c(x) + i\theta\chi(x) - i\bar{\theta}\bar{\chi}(x) + \frac{1}{2}\theta^2 [M(x) + iN(x)] \\ - \frac{1}{2}\bar{\theta}^2 [M(x) - iN(x)] - \theta\sigma^m\bar{\theta}A_m(x) \\ + i\theta^2\bar{\theta} \left[\bar{\lambda}(x) + \frac{1}{2}\bar{\sigma}^m\partial_m\chi(x) \right] \\ - i\bar{\theta}^2\theta \left[\lambda(x) + \frac{i}{2}\sigma^m\partial_m\bar{\chi}(x) \right] \\ + \frac{1}{2}\theta^2\bar{\theta}^2 \left[D(x) + \frac{1}{2}\square c(x) \right]$$

now note if we take a chiral superfield and add it to an anti-chiral superfield, we get:

$$(3.116) \quad \Phi + \bar{\Phi} = (\varphi + \bar{\varphi}) + \sqrt{2} \left(\theta\psi + \bar{\theta}\bar{\psi} + \theta^2 F + \bar{\theta}^2 \bar{F} \right) + i\theta\sigma^m\bar{\theta}\partial_m(\varphi - \bar{\varphi}) \\ + \frac{i}{\sqrt{2}}\theta^2\bar{\theta}\bar{\sigma}^m\partial_m\psi + \frac{i}{2}\bar{\theta}^2\theta\sigma^m\partial_m\bar{\psi} + \frac{1}{4}\theta^2\bar{\theta}^2\square(\varphi + \bar{\varphi})$$

Proposition 3.1.

$$(3.117) \quad \delta V = \Phi + \bar{\Phi}$$

as the superfield gauge transformation for V .

Note that this means the component transformations are algebraic for C , χ , M , N . The gauge where these all vanish is called the Wess-Zumino (WZ) gauge. This gauge actually breaks supersymmetry. The $N = 1$ vector multiplet is given by (A_m, λ, D) . Note that in WZ gauge, $V^3 = 0$. Now what we would like, is to find a superfield invariant under this gauge transformation. To do this, we make use of a chiral “projector.” We can define the field strength superfield:

$$(3.118) \quad W_\alpha = -\frac{1}{4}\bar{D}^2 D_\alpha V$$

this kills both chiral and anti-chiral fields, and is itself chiral. This invariance is given by:

$$(3.119) \quad \delta W_\alpha = -\frac{1}{4}\bar{D}^2 D_\alpha (\Phi + \bar{\Phi}) = -\frac{1}{4}\bar{D}^2 D_\alpha \Phi$$

$$(3.120) \quad = \frac{1}{4}\bar{D}^{\dot{\alpha}} \{\bar{D}_{\dot{\alpha}}, D_\alpha\} \Phi = -\frac{i}{2}\bar{D}^{\dot{\alpha}} \partial_{\alpha, \dot{\alpha}} \Phi = 0$$

This satisfies what are sometimes called “Bianchi identities” given by:

$$(3.121) \quad \bar{D}_{\dot{\alpha}} W_\alpha = 0 \quad D^\alpha W_\alpha = \bar{D}_{\dot{\alpha}} \bar{W}^{\dot{\alpha}}$$

the first one just says chirality, the second says D is real. Now looking at the components, we do in fact find the field strength for A :

$$(3.122) \quad W_\alpha = -i\lambda_\alpha + \left[\delta_\alpha^\beta D - \frac{i}{2} (\sigma^m \bar{\sigma}^n)^\beta_\alpha (\partial_m A_n - \partial_n A_m) \right] \theta_\beta + \theta^2 \partial_{\alpha\dot{\alpha}} \bar{\lambda}^{\dot{\alpha}}$$

Now we want to find an action. The idea here is that every top component $D^2 \bar{D}^2 \Psi$ or $D^2 W_\alpha$ of a superfield is a total derivative. The notation we will use will be:

$$(3.123) \quad \int d^2\theta S(x, \theta, \bar{\theta}) \equiv \frac{1}{16} D^2 \bar{D}^2 S|$$

$$(3.124) \quad \int d^2\theta S'(x, \theta, \bar{\theta}) = -\frac{1}{4} D^2 S'|$$

which are supersymmetry invariant. For us, our action is given by:

$$(3.125) \quad \mathcal{L} = \int d^2\theta W^\alpha W_\alpha + \int d^2\bar{\theta} \bar{W}_{\dot{\alpha}} \bar{W}^{\dot{\alpha}}$$

$$(3.126) \quad \sim W^\alpha D^2 W_\alpha| + \dots \sim \lambda \partial \bar{\lambda} + \dots + F^{mn} F_{mn} + iF \wedge F$$

To get the equations of motion, vary V to get $D^\alpha W_\alpha = 0$. This leads to

$$(3.127) \quad D^2 W_\alpha| = (\partial \bar{\lambda})_\alpha = 0$$

So we really don't have to do anything with components if we don't want to.

5.1.1. *Matter.* We start with just a U(1) gauge theory. We attempt to introduce l species of matter given by Φ_l . For this, these chiral superfields will have the transformations:

$$(3.128) \quad \Phi'_l = e^{-it_l \lambda} \Phi_l$$

We have kinetic terms:

$$(3.129) \quad K[\Phi, \bar{\Phi}] = \int d^4\theta \bar{\Phi}_l \Phi_l$$

This is sometimes called the Kähler potential. We also have the potential terms:

$$(3.130) \quad W[\Phi] = \int d^2\theta [m_{ij} \Phi_i \Phi_j + g_{ijk} \Phi_i \Phi_j \Phi_k]$$

this is the most general renormalizable U(1) invariant coupling setup.

In the local setting, we take Λ to be spacetime dependent, and we take the following transformations:

$$(3.131) \quad \Phi'_l = e^{-it_l \Lambda} \Phi_l$$

This does not affect the potential term, but the kinetic term is changed. We now have:

$$(3.132) \quad \bar{\Phi}'_l \Phi'_l = \bar{\Phi}_l \Phi_l e^{it_l (\bar{\Lambda} - \Lambda)}$$

but we can simple compensate this with $\delta V = i(\Lambda - \bar{\Lambda})$. So now the Kähler potential is modified to:

$$(3.133) \quad K[\Phi, \bar{\Phi}] = \int d^4\theta \bar{\Phi}_l e^{t_l V} \Phi_l$$

In the nonabelian setting, for $\Lambda_{ij} = T_{ij}^a \Lambda_a$ we get

$$(3.134) \quad e^{V'} = e^{-i\bar{\Lambda}} e^V e^{i\Lambda}$$

$$(3.135) \quad W_\alpha = -\frac{1}{4} \bar{D}^2 e^{-V} D_\alpha e^V \quad \delta W_\alpha = e^{-i\Lambda} W_\alpha e^{i\Lambda}$$

Now one more invariant is called the Fayet-Iliopoulos term:

$$(3.136) \quad \xi \int d^4\theta V$$

which is relevant for supersymmetry breaking.

5.2. Yang-Mills with $N = 2$ SUSY. In this section we consider $N = 2$ SUSY working with $N = 1$ superfields as sort of components. Recall the general $N = 1$ action was given by:

$$(3.137) \quad \mathcal{L} = \int d^2\theta W^\alpha W_\alpha + \int d^4\theta \bar{\Phi} e^V \Phi + \int d^2\theta W[\Phi] + \xi \int d^4\theta V$$

In order to enhance this to the $N = 2$ case we take

$$(3.138) \quad \mathcal{L} = \frac{1}{16\pi} \text{Im} \left[\tau \int d^2\theta W^\alpha W_\alpha + \tau \int d^4\theta \bar{\Phi} e^{-2gV} \Phi \right]$$

where

$$(3.139) \quad \tau = \theta/2\pi + 4\pi i/g^2$$

In components this is written:

$$(3.140) \quad \mathcal{L} = -\frac{1}{4}F_{mn}F^{mn} - i\lambda\sigma^m\nabla_m\bar{\lambda} + \frac{\theta}{32\pi^2}F_{mn}\tilde{F}^{mn} + |\nabla_m\varphi|^2 \\ - i\bar{\Psi}\bar{\sigma}^m\nabla_m\psi + \bar{F}F - g\bar{\varphi}[D, \varphi] - \sqrt{2}ig\bar{\varphi}\{\lambda, \psi\} + \sqrt{2}ig\bar{\varphi}[\bar{\lambda}, \varphi]$$

Connection to top YM: For $N = 2$ we have

$$(3.141) \quad \text{SU}(2)_L \times \text{SU}(2)_R \times \underbrace{\text{SU}(2)_A \times \text{U}(1)_R}_{R\text{-symmetry}}$$

for the twisted theory, take $\text{SU}(1)'_R$ as the diagonal subgroup of $\text{SU}(2)_R \times \text{SU}(2)_A$.

In general, for “manifest” $N = 1$ SUSY, we can introduce a second set of SUSY generators and corresponding $\tilde{\theta}_\alpha$ and $\bar{\tilde{\theta}}_{\dot{\alpha}}$. Now we want the $N = 2$ chiral superfield. Here this means something in a coordinate basis independent of $\bar{\theta}$ and $\bar{\tilde{\theta}}$.

$$(3.142) \quad \Psi = \Phi + \tilde{\theta}^\alpha W_\alpha + \bar{\tilde{\theta}}^2 G(\Phi, W)$$

where G is just some unimportant function of φ and W . The simplest action here for Φ is given by:

$$(3.143) \quad \text{Im} \int d^2\theta d^2\tilde{\theta} \Phi^2$$

then the most general one is given by

$$(3.144) \quad \text{Im} \int d^2\theta d^2\tilde{\theta} \mathcal{F}(\Psi)$$

\mathcal{F} is called the $N = 2$ pre-potential^{3.2} The importance here is that our simple kinetic form could potentially hold this general form after renormalization. Classically,

$$(3.145) \quad \mathcal{F}_\ell(\Psi) = \frac{1}{2} \text{tr} [\tau\Psi]$$

In $N = 1$ language,

$$(3.146) \quad \mathcal{L} = \frac{1}{16\pi} \text{Im} \left[d^2\theta \mathcal{F}''(\Phi) W^\alpha W_\alpha + \int d^4\theta \bar{\Phi} e^{-2gV} \mathcal{F}'(\Phi) \right]$$

5.3. Seiberg-Witten theory. This is looking at the low energy effective action of the $N = 2$ theory as detailed in the original paper by Seiberg and Witten [21]. As it turns out, this gives a lot of exact results. Why is this of interest to us? As it turns out, $N = 2$ YM with SUSY is a “sweet-spot” between $N = 1$ which is not exactly solvable and $N = 4$ which is exactly solvable classically, that is it is a finite theory. This is also a very nice realization of the so called S -duality from string theory. This is also called the electric-magnetic duality, and the Montonen-Olive duality. This is important because it relates strongly coupled theories to weakly coupled theories. For example, in QCD we have a large coupling constant, so perturbative calculations don’t hold much water. This stands in comparison with QED where we have a small coupling. There is also an example of confinement in this theory. A lot of things in this low energy regime have their origins in string theory from brane constructions. As with most papers by Witten, there are many interesting mathematical results, in this case there is attention given to the study of elliptic curves.

^{3.2} There is a formal notion of pre-potential, and this is in fact not a pre-potential in this sense.

Consider a gauge group given by $SU(2)$. We do the natural thing, and first consider the classical potential. The auxiliary part of this action is given by:

$$(3.147) \quad \mathcal{L}_{\text{aux}} = \frac{1}{g^2} \left[\frac{1}{2} D^2 + \bar{F}F - g\varphi [D, \varphi] \right]$$

now integrating out F and D gives:

$$(3.148) \quad \mathcal{L}_{\text{aux}} = -\frac{1}{2} ([\bar{\varphi}, \varphi])^2 \rightsquigarrow V(\varphi) = \frac{1}{2} \text{tr}([\bar{\varphi}, \varphi])^2$$

For a vacuum φ_0 , $N = 2$ supersymmetry is unbroken when $V(\varphi_0) = 0$.

Note this is for a pure YM theory. Even though it looks like we have coupled additional matter, this is really just pure $N = 2$ YM. Of course if we add mass here, we will get additional contributions to the potential. That is we don't consider hypermultiplets at this point. So we are working on the Coulomb branch of the moduli space. The moduli space is just the space of gauge inequivalent vacua.

We can decompose

$$(3.149) \quad \varphi(x) = \sum_{i=1}^3 (a_i(x) + ib_i(x)) \sigma_i$$

now $a_1 = a_2 = 0$ by gauge transformations, and $b_1 = b_2 = 0$ by the condition $[\bar{\varphi}, \varphi] = 0$. Then we can take $a \equiv a_3 + ib_3$ and $\varphi = \frac{1}{2} a \sigma_3$ which leads to an invariant $\text{tr} \varphi^2$. Then take the coordinate on the moduli space μ to be $u = \langle \text{tr} \varphi^2 \rangle$. $\langle \varphi \rangle = \frac{1}{2} a \sigma_3$.

REMARK 3.1. Note that since $\langle \varphi \rangle \neq 0$ the $SU(2)$ is broken to a $\mathcal{U}(1)$ by a Higgs mechanism which gives mass to some of our fields.

In $N = 1$ language we have the component expression:

$$(3.150) \quad \mathcal{L} = \frac{1}{4\pi} \text{Im} \left[\mathcal{F}''(\varphi) (\partial_\mu \varphi)^2 + \dots \right]$$

this suggests that \mathcal{F}'' plays the role of a σ -model metric on our moduli space. More precisely, this plays the role of a Kähler metric on a Kähler manifold. On such a manifold we have a real Kähler potential defined by

$$(3.151) \quad g_{i\bar{j}} = \frac{\partial^2 K(z, \bar{z})}{\partial z^i \partial \bar{z}^{\bar{j}}} \rightsquigarrow ds^2 = g_{i\bar{j}} dz^i d\bar{z}^{\bar{j}}$$

for us we will simply have

$$(3.152) \quad k(\Phi, \bar{\Phi}) = \text{Im}(\bar{\Phi} \mathcal{F}'(\Phi))$$

which gives us the line element:

$$(3.153) \quad ds^2 = \text{Im} \mathcal{F}''(a) da d\bar{a} = \text{Im} \tau(a) da d\bar{a}$$

But now we have a problem. It's nice that this \mathcal{F}'' is holomorphic, but it is actually not positive definite. The solution is that the moduli space actually has singularities, so we need to find coordinate patches to construct a metric which is positive definite in this geometrically local sense. Note that in [21] there is an exact non-perturbative result obtained for $\mathcal{F}(\Psi)$ given by:

$$(3.154) \quad \mathcal{F}_{1\text{-loop}}(\Psi) = \frac{1}{2\pi} \Psi^2 \log \frac{\Psi^2}{\Lambda^2} \quad \Delta \mathcal{F}_{\text{non-pert}} = \sum_{k=1}^{\infty} \mathcal{F}_k \left(\frac{\Lambda}{\Psi} \right)^{4k} \Psi^2$$

We now perform the Legendre transformation:

$$(3.155) \quad \mathcal{F}_D(\Phi_D) = \mathcal{F}(\Phi) - \Phi\Phi_D$$

so we take Φ_D to be \mathcal{F}' , then this means $\mathcal{F}'_D(\Phi) = -\Phi$. As it turns out, this doesn't affect the Kähler potential at all, but it has a very important effect on the W^2 part.

$$(3.156) \quad \int \mathcal{D}V \exp \left[\frac{1}{16\pi} \int d^4x d^2\theta \operatorname{Im} [\mathcal{F}''(\Phi) W^\alpha W_\alpha] \right] \\ \cong \int \mathcal{D}W \mathcal{D}V_D \exp \left[\frac{i}{16\pi} \operatorname{Im} \int d^4x [d^2\theta \mathcal{F}''(\Phi) W^2 + 2 d^4\theta V_D D^\alpha W_\alpha] \right]$$

where we have introduced a suggestively named Lagrange multiplier V_D . Now we can integrate out W in favor of its Lagrange multiplier. A simple calculation shows:

$$(3.157) \quad \int d^4\theta V_D D^\alpha W_\alpha = - \int d^2\theta W_D^\alpha W_\alpha \quad W_{\alpha D} = -\frac{1}{4} \bar{D}^2 D_\alpha V_D$$

so integrating out $W_\alpha = W_D/\mathcal{F}(\Phi)$,

$$(3.158) \quad \int \mathcal{D}V_D \exp \left[\frac{i}{16\pi} \operatorname{Im} \int d^4x d^2\theta \left(-\frac{1}{\mathcal{F}''} W_D^\alpha W_{D\alpha} \right) \right]$$

so this is the S -dual formulation of the theory. In terms of the coordinates a, \bar{a} on the moduli space we have that $a_D = \mathcal{F}'(a)$ gives access to the remainder of the moduli space.

6. Chern-Simons theory

This is a 3-dimensional gauge theory given by a 1-form gauge field A . In components this is written $A_\mu dx^\mu$. Now for a gauge group G , this is given by:

$$(3.159) \quad A = A_\mu^a T^a dx^\mu$$

which is Lie algebra valued. T^a are the generators and satisfy

$$(3.160) \quad [T^a, T^b] = i f^{ab}_c T^c$$

Then we can raise and lower indices using:

$$(3.161) \quad h^{ab} = \operatorname{Tr}(T^a T^b)$$

The Chern-Simons action we will consider is given by:

$$(3.162) \quad S_{\text{cs}} = \frac{k}{2\pi} \int \operatorname{tr} \left(A \wedge dA + \frac{2}{3} A \wedge A \wedge A \right)$$

$$(3.163) \quad = \frac{k}{4\pi} \int \epsilon^{\mu\nu\sigma} \left(A_{\mu a} \partial_\nu A_\sigma^a + \frac{2}{3} f_{abc} A_\mu^a A_\nu^a A_\sigma^a \right) d^3x$$

6.1. Quantum hall effect. In 3-dimensions, there is something special about YM theory. Recall the YM action is given by:

$$(3.164) \quad S_{\text{ym}} = \frac{1}{4g^2} \int \operatorname{Tr}(F \wedge \star F)$$

REMARK 3.2. If we have

$$(3.165) \quad F = \frac{1}{2} F_{\mu\nu} dx^\mu \wedge dx^\nu$$

then

$$(3.166) \quad \star F = \frac{1}{4} \epsilon_{\mu\nu}{}^{\sigma\tau} F_{\sigma\tau} dx^\mu \wedge dx^\nu$$

and we have

$$(3.167) \quad F \wedge \star F = \text{tr} (F_{\mu\nu} F^{\mu\nu}) \frac{1}{4!} \epsilon_{\mu\nu\sigma\tau} dx^\mu \wedge dx^\nu \wedge dx^\sigma \wedge dx^\tau$$

So we take the total action to be:

$$(3.168) \quad S_{\text{Tot}} = S_{\text{ym}} + S_{\text{cs}}$$

What does this theory look like at low energy? From “dimensional analysis” we get that

$$(3.169) \quad [g^2] = \begin{cases} 0 & 4D \\ [M] & 3D \end{cases}$$

the first case is referred to as the running of the coupling constant. The second case follows from the fact that A_μ also has the dimension of mass. In any case, this means the coupling g becomes large in this low energy limit, so we can disregard the YM action, and concern ourselves with the Chern-Simons action.

We now consider the quantum hall effect (QHE). There are two versions of this. The first is the integer case, which is completely understood and the second is the fractional case which is largely an open problem. Let’s first consider the integer case. Consider a thin film, modelled as a 2-dimensional strip of width L under the influence of a very strong magnetic field of strength B . Then apply a voltage V_y over the width of the strip, and measure the current I_x along the strip. Recall that the resistance is given by $\rho_H = V_y/I_x$. Now if the charge carriers have charge Q and density n , we have

$$(3.170) \quad \rho_H = \frac{V_y}{I_x} = \frac{E_y L}{J_x L} = \frac{E_y}{J_x} \quad J_x = nqv_x$$

then we have $E_y = v_x B$ and can write

$$(3.171) \quad \rho_H = \frac{E_y}{nqv_x} = \frac{B}{nq}$$

Note that more generally,

$$(3.172) \quad \sum_{j \in \{x,y\}} \rho_{ij} J_j = E_i$$

and if we measure ρ_{xy} in terms of B , we get a linear increase which plateaus at values $\rho_{xy} = h/(\nu e^2)$ for $\nu \in \mathbb{Z}$. If we measure ρ_{xx} , we get bumps in the regions where ρ_{xy} is linear, but it drops to 0 at the plateaus.

Why do we see this behaviour? Let’s consider a single particle with charge q in $\vec{B} = B\hat{z}$. Then we have the lowest two Landau levels with an energy gap given by $\hbar\omega_B = qB/m$. The degeneracy is

$$(3.173) \quad \varphi_0 = \frac{2\pi\hbar}{|e|}$$

so the number of states is given by $N = BA/\varphi_0$. So each state has surrounded itself with this magnetic flux, and the question is how many we have. Now these values $1/\nu$ occur

exactly when an integer number of Landau levels are filled. So ν gives the number of LL filled. But this would lead us to think the plateau would only occur at a single point. As it turns out, the impurities in the material cause these larger plateaus to form. Now if ν states are filled, then

$$(3.174) \quad n = \frac{\nu N}{A} = \frac{\nu B}{\varphi_0}$$

so for $q = -c$,

$$(3.175) \quad \rho_H = \frac{1}{\nu} \left(\frac{2\pi\hbar}{c^2} \right)$$

which explains the ρ_{xy} behavior. The ρ_{xx} behaviour is determined by the question of determining J_x solely from turning on $E_x = \rho_{xx}J_x + \rho_{xy}J_y$. Since $\rho = \sigma^{-1}$, if we can show that $\sigma_{xx} = 0$, then σ would have zeros on the diagonal which would mean $\rho_{xx} = 0$. Now we can make a Galilean boost which brings E to zero. In particular, with this large B we can find such a velocity where there will not be any current. But when we boost back to our reference frame, we can only get a y current, not an x -current, since we boosted in the y direction. But this only applies to pure materials, since the impurities boost as well, so we do indeed get this ρ_{xx} behaviour.

The fractional quantum hall effect is similar to the integer case, only now we get plateaus at fractional values of ν . For example, we might have a plateau at $\nu = 2/3$ and $\nu = 3/5$. As it turns out, the numbers that appear here are called filling fractions. Somore additional examples are given by:

$$(3.176) \quad \frac{1}{3} \quad \frac{2}{5} \quad \frac{3}{7} \quad \frac{2}{3} \quad \frac{3}{5} \quad \frac{4}{7} \quad \dots$$

or more generally:

$$(3.177) \quad \nu = \frac{p}{q} = \frac{n}{2mn \pm 1}$$

This problem is more difficult because we can't think about this problem as free electrons in a field. Here this interpretation means Landau levels are partially filled. So if N is the number of states in LL , and $\nu = 1/3$, then there are

$$(3.178) \quad \begin{pmatrix} N \\ N/3 \end{pmatrix}$$

ways to fill the levels which is very large.

Now what does field theory have to say about this? The field theory approach takes A_{ext} to be the EM potential which creates the B -field $\vec{B} = B\hat{z}$ as well as \vec{E} which leads to V_y . Then our action will be a function of A_{ext} and the electrons in the sample which is given by $\hat{\varphi}(\vec{r}, t)$. Now we attempt the path integral

$$(3.179) \quad \exp\left(-\frac{i}{\hbar} S_{\text{eff}}[A_{ext}]\right) = \int \exp\left(-\frac{i}{\hbar} S[A_{ext}, \Psi][\mathcal{D}\Psi]\right)$$

The claim is that we have:

$$(3.180) \quad S_{\text{eff}} = \frac{k}{4\pi} \int A \wedge dA$$

which is Lorentz invariant and topological since it is independent of the metric, and has no propagating degrees of freedom. This is called the abelian Chern-Simons action. Now for

$J_i = \sigma_{ij} E_j$, we seek:

$$(3.181) \quad \langle J_i(x) \rangle_{A_{ext}} = \left\langle \frac{\delta S}{\delta A_{ext}} \right\rangle = \int \left(\frac{\delta S}{\delta A} \right) e^{iS/\hbar} \mathcal{D}\Psi$$

$$(3.182) \quad = \frac{\hbar}{i} \frac{\delta}{\delta A} \int e^{iS/\hbar} \mathcal{D}\Psi = \frac{\delta}{\delta A} \left(e^{iS_{\text{eff}}/\hbar} \right)$$

$$(3.183) \quad = \frac{\delta S_{\text{eff}}}{\delta A} = \frac{k}{2\pi} \epsilon_{ijl} F_{jl}$$

so macroscopically, we get the relation:

$$(3.184) \quad J_i = \frac{k}{2\pi} \epsilon_{ij} E_j$$

That is, for $\sigma_{xy} = k/2\pi$ we get $\rho_{xy} = 2\pi/k$.

We can also obtain the density of states in a similar manner:

$$(3.185) \quad \langle J^0 \rangle = \frac{k}{2\pi} B$$

which should be compared with

$$(3.186) \quad n = \frac{\nu e B}{2\pi k} = \frac{\nu B}{\varphi_0}$$

So how do we know k is an integer? Well if we are just on \mathbb{R}^2 there are no restrictions on k . But if we put the theory on some compact space with some genus, such as a sphere, torus, or general Riemann surface we get some restrictions. We can also put the theory on some surface with boundary to get some restrictions on our k value.

First we put the theory on the torus T^2 , which we think of as being a periodic space with coordinates x_1 and x_2 . The action is of the form

$$(3.187) \quad S = \frac{k}{4\pi} \int A \wedge dA$$

We now quantize this theory. First take the $A_0 = 0$ gauge fixing condition. But recall we still have the condition:

$$(3.188) \quad 0 = \frac{\delta A}{\delta A_0} = F_{12}$$

so there is in fact no field strength. Now this condition along with gauge invariance allows us to set the gauge field to be of the form:

$$(3.189) \quad A = a_1 dx_1 + a_2 dx_2$$

where a_1 and a_2 are just functions of time. This means we have

$$(3.190) \quad S = \frac{k}{4\pi} \int A \wedge dA = \frac{k}{4\pi} \int (a_2 \dot{a}_1 - a_1 \dot{a}_2) dt$$

$$(3.191) \quad = \frac{k}{2\pi} \int a_2 \dot{a}_1 dt dx dy + \text{boundary terms}$$

Now we have the canonical relation:

$$(3.192) \quad [a_2, a_1] = -\frac{2\pi i}{k}$$

But what is the Hilbert space? Recall the residual gauge transformation

$$(3.193) \quad \Lambda = 2\pi N dx_1$$

where $N \in \mathbb{Z}$. This doesn't look like a gauge transformation though, since this is not single valued. But this is ok, because only $e^{i\Lambda} \in U(1)$ actually needs to be single valued. So let's see how this acts on A :

$$(3.194) \quad \begin{aligned} A &\xrightarrow{\Lambda} a_1 dx_1 + a_2 dx_2 + 2\pi N dx_1 \\ a_1 &\xrightarrow{2\pi N x_2} a_1 + 2\pi N \\ a_2 &\xrightarrow{2\pi M x_1} a_2 + 2\pi M \end{aligned}$$

This means

$$(3.195) \quad a_1 \sim a_1 + 2\pi \qquad a_2 \sim a_2 + 2\pi$$

a_1 and a_2 are not really what we want. So instead, we define:

$$(3.196) \quad W_1 = e^{ia_1} \qquad W_2 = e^{ia_2}$$

Now take C_1 and C_2 to be loops along the two copies of S^1 making up T^2 . Then we have that

$$(3.197) \quad E_i = \exp \left(i \oint_{C_i} A \right)$$

If we were to just integrate A around these loops we would end up with an extra shift, whereas if we integrate the exponentials these constants are in the exponential. Now we can write:

$$(3.198) \quad W_1 W_2 W_1^{-1} W_2^{-1} = e^{-[a_1, a_2]} = e^{-2\pi i/k} = -\omega$$

where ω denotes a k th root of unity. In particular,

$$(3.199) \quad W_2^{-1} W_1 W_2 = \omega^{-1} W_1$$

So we can write:

$$(3.200) \quad W_1 = e^{i\varphi} \begin{pmatrix} 1 & \cdots & \cdots & \cdots \\ \cdots & \omega & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \omega^{k-1} \end{pmatrix}$$

$$(3.201) \quad W_2 = e^{i\varphi'} \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & \cdots \\ 0 & 0 & 1 & 0 & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 0 & 0 & \cdots & \cdots & 0 \end{pmatrix}$$

where φ, φ' are effectively arbitrary. Now if we want a finite dimensional Hilbert space, then k must be an integer.

We can also geometrically quantize our phase space T^2 . Take

$$(3.202) \quad u = a_1 + ia_2 \qquad \bar{u} = a_1 - ia_2 \qquad [u, \bar{u}] = \frac{2\pi}{k}$$

so if we have $u^\dagger = \bar{u}$ this leads to $4\pi\partial_u/k$. But what are the wavefunctions? Well we can't take

$$(3.203) \quad \psi(a_1, a_2) \rightarrow \psi(u, \bar{u})$$

so we have to pick one. This means

$$(3.204) \quad \frac{\partial}{\partial \bar{u}} \psi = 0$$

which means these are holomorphic. Which means

$$(3.205) \quad u^\dagger = \frac{4\pi}{k} \frac{d}{du}$$

But how could this be? Let's see what this means from a different perspective. Take the inner product:

$$(3.206) \quad \langle \chi | \psi \rangle := \int \chi^* \psi d^2 u$$

which means:

$$(3.207) \quad \left\langle \chi \left| \frac{d}{du} \psi \right. \right\rangle = \int \chi^* \frac{d}{du} \psi d^2 u = - \int \frac{d\chi^*}{du} \psi d^2 u + \frac{k}{4\pi} \int \bar{u} \chi^* \psi d^2 u$$

$$(3.208) \quad = \frac{k}{4\pi} \int \bar{u} \chi^* \psi d^2 u$$

so the fix here is to instead define the inner product to be:

$$(3.209) \quad \langle \chi | \psi \rangle := \int_{T^2} \chi(u)^* \psi(u) e^{-k|u|^2/(4\pi)} d^2 u$$

which means

$$(3.210) \quad \left\langle \chi \left| \frac{d}{du} \psi \right. \right\rangle = \frac{k}{4\pi} \int \bar{u} \chi^* \psi e^{-k|u|^2/(4\pi)} d^2 u$$

So the problem is that

$$(3.211) \quad \chi(u)^* \psi(u) e^{-k|u|^2/(4\pi)} = \chi(u+2\pi)^* \psi(u+2\pi) e^{-k|u+2\pi|^2/4\pi}$$

$$(3.212) \quad = \chi(u+2\pi i)^* \psi(u+2\pi i) e^{-k|u+2\pi i|^2/4\pi}$$

$$(3.213)$$

So the solution is to demand:

$$(3.214) \quad \psi(u+2\pi) = \psi(u) \exp(k(u+\pi)/(2\pi))$$

$$(3.215) \quad \psi(u+2\pi i) = \psi(u) \exp(k(-iu+\pi)/(2\pi))$$

So what is the Hilbert space of holomorphic function which satisfy eq. (3.215)? Well if $k \notin \mathbb{Z}$, there are no solutions. But if $k \in \mathbb{Z}^+$, then this just means

$$(3.216) \quad \psi = \sum_{j=0}^{k-1} c_j \psi_j$$

and the explicit solutions are given by:

$$(3.217) \quad \psi_j(u) = e^{ku^2/(8\pi)} \sum_{n=-\infty}^{\infty} e^{(kn+j)^2/(2k)} \exp(i(kn+j)u)$$

for $j \in \{1, \dots, k-1\}$ and $q = e^{-2\pi}$.

We now consider a generalization of the previous considerations. Consider a general compact phase space, and now quantize the theory. Equip this space with a symplectic 2-form ω . This means

$$(3.218) \quad \omega - \sum_{i=1}^m dx_i \wedge dp_i \quad d\omega = 0$$

So $\omega^{\mu\nu}$ is $2m$ dimensional. Since this is Symplectic we have:

$$(3.219) \quad \omega_{\mu\nu} \rightarrow (\omega^{-1})^{\mu\nu} \quad (\omega^{-1})^{\mu\nu} \omega_{\nu\sigma} = \delta^\mu_\sigma$$

EXAMPLE 3.8. In the case of a sphere,

$$(3.220) \quad \omega = \sin \theta d\theta \exp d\rho$$

Taking local coordinates x^i we have

$$(3.221) \quad [x^l, x^j] = -i\omega^l_j$$

We now equip this with a complex structure. Pair our local coordinates into z^α for $\alpha \in \{1, \dots, m\}$. Then take

$$(3.222) \quad \omega_{\alpha\bar{\beta}} = -\omega_{\bar{\beta}\alpha} \quad \omega_{\alpha\beta} = \omega_{\bar{\alpha}\bar{\beta}} = 0$$

Now locally we have

$$(3.223) \quad \omega_{\alpha\bar{\beta}} = i\partial_\alpha \partial_{\bar{\beta}} K$$

where K is the Kähler function which means

$$(3.224) \quad [\partial_{\bar{\alpha}} K, z^{\bar{l}}] = \delta_{\bar{\alpha}}^{\bar{l}}$$

EXAMPLE 3.9. On T^2 , $Z^1 = u$ and $Z^{\bar{1}} = \bar{u}$. Then $K = \frac{k}{4\pi} \bar{u}^2$.

So now define the inner product to be:

$$(3.225) \quad \langle \chi | \psi \rangle = \int \chi^* \psi e^{-K} d^{2M} z$$

so now our wave functions are

$$(3.226) \quad \psi(z^1, \dots, z^M)$$

But these Kähler functions are only defined locally, so for two patches we might have two different functions K_I and K_{II} . Then we have

$$(3.227) \quad K_{II} = K_I + f + \bar{f}$$

where f is holomorphic. In particular,

$$(3.228) \quad f = \frac{k}{4\pi} u + \frac{k}{8\pi}$$

and

$$(3.229) \quad \psi_I = e^f \psi_{II}$$

So the problem is to find holomorphic sections ψ_I, Ψ_I, \dots such that

$$(3.230) \quad \psi_I = e^{f_{I,II}} \psi_{II}$$

so we are looking for a holomorphic line bundle over our phase space. This is just a space of sections.

FQHE: Take $\nu \in \mathbb{Z}$. Then sending $A \rightarrow A + d\Lambda$ we get

$$(3.231) \quad \frac{k}{4\pi} \int_M A \wedge dA \rightarrow \frac{k}{4\pi} \int_M A \wedge dA + \frac{k}{4\pi} \int_{\partial M} \Lambda dA$$

which is not gauge invariant. Now consider this on a cylinder with t along the middle axis, and θ the angular coordinate. Then $dA = F$ and

$$(3.232) \quad \int F_{t\theta} = dt d\theta = 2\pi m$$

$$(3.233) \quad \oint A_\theta d\theta \Big|_{t_i}^{t_f} = 2\pi m$$

so according to this formalism, there is no such FQHE.

But in fact, there can be additional topological degrees of freedom, which means we can have fractionally charged particles on the boundary which will give us the FQHE. We know we have the coupling $J \wedge A$ where $d \star A = 0$ and $\star J = da/2\pi$. So the full action is

$$(3.234) \quad \int \star J \wedge A + \frac{\tilde{k}}{4\pi} \int d \wedge da = \frac{1}{2\pi} \int da \wedge A + \frac{k}{4\pi} \int a \wedge da$$

where we have the equation of motion:

$$(3.235) \quad [a] = 0 = \tilde{k} da + da$$

which leads to:

$$(3.236) \quad J = -\frac{1}{2\pi\tilde{k}} \star dA$$

which of course gives us our fractional case.

6.2. Non-abelian Chern-Simons. Consider the action

$$(3.237) \quad S = \frac{k}{4\pi} \int \text{tr} \left(A \wedge dA + \frac{2}{3} A \wedge A \wedge A \right)$$

now one way to quantize is to take

$$(3.238) \quad A_0 = 0 \implies F = 0$$

now what about some compact Riemann surface? When we consider gauge fields on this surface, there is no field strength, which seems to mean there is nothing going on. But we can still have a nontrivial gauge field which satisfies

$$(3.239) \quad 0 = dA + A \wedge A$$

and if we look at these fields modulo

$$(3.240) \quad A \sim \Omega^{-1} A \Omega + \Omega^{-1} d\Omega$$

this gives a finite moduli space of what are called flat connections. Since this is finite, we can define

$$(3.241) \quad \int \delta A_z \wedge \delta A_{\bar{z}} = \omega_{\alpha\bar{\beta}} \delta z^\alpha \alpha z^{\bar{\beta}}$$

then we can do geometric quantization, and get a finite dimensional Hilbert space. What is this good for? Jones polynomials:

If we have

$$(3.242) \quad O(x) = \text{tr} F_{\mu\nu} F^{\mu\nu}$$

then

$$(3.243) \quad \langle O(x_1) O(x_2) \cdots \rangle = 0$$

But now if we take some knot C we can calculate the more general

$$(3.244) \quad \langle W(C) \rangle$$

which gives $f(K)$ which leads to the knot invariant

$$(3.245) \quad q = \exp\left(\frac{2\pi i}{k+2}\right)$$

See [26] for a more in depth approach on this subject.

7. BRST-BV Formalism

7.1. Motivation. This is a mix of the tools from BRST formalism, such as the cohomological charge, $Q^2 = 0$, along with new tools. In addition to the fields (ghosts anti-ghosts, ghosts for ghosts) we will also introduce antifields. This is why this formalism is also called anti-field BRST. We will justify the addition of these anti-fields by defining what it means for a symmetry algebra to be open. We will also see an anti-bracket, (\cdot, \cdot) which is in some sense a generalization of a Poisson bracket. At the quantum level, we will also introduce an operator Δ which is a second order operator.

The added BV stands for Batalin-Vilovinsky. They did the majority of their work with this formalism between 1981-1985. This formalism will be important for the following applications:

- (1) Some theories cannot be quantized in a straight-forward way using the FP procedure. In particular this is the case when the symmetry object is not closed. This just means that instead of $[\delta_\alpha, \delta_\beta] = f_{\alpha\beta}{}^\gamma \delta_\gamma$ we have

$$(3.246) \quad [\delta_\alpha, \delta_\beta] = f_{\alpha\beta}{}^\gamma(\varphi) \delta_\gamma + \text{EoM}(\varphi)$$

- (2) Even for YM in, say, $3+1$ dimensions, when we try to show renormalizability to all loop orders, BRST is not enough. This was shown using the BV formalism by Zinn-Justin in [28].
- (3) This is important for understanding the structure of both global and gauge anomalies. The important question here is determining when a symmetry is necessarily anomalous in quantum mechanics. This is cohomological in nature, and the BRST formalism is sufficient for this. But in order to show:

Proposition 3.2. *The chiral anomaly (in the standard model) is one-loop exact.*

we need the full BRST-BV extension of the Q -cohomology.

7.2. Open symmetry algebras. Start with the action:

$$(3.247) \quad S_0[\varphi] = \int d^D x \mathcal{L}_0(\varphi^i, \partial_\mu \varphi^i, \dots, \partial_{\mu_1 \dots \mu_k} \varphi^i)$$

The equation of motion is given by:

$$(3.248) \quad \frac{\delta S_0}{\delta \varphi^i(x)} \equiv \frac{\delta \mathcal{L}_0}{\delta \varphi^i}(x) = 0$$

where we recall

$$(3.249) \quad \frac{\delta \mathcal{L}_0}{\delta \varphi^i}(x) = \frac{\partial \mathcal{L}_0}{\partial \varphi^i} - \partial_\mu \frac{\partial \mathcal{L}_0}{\partial (\partial_\mu \varphi^i)} + \partial_\mu \partial_\sigma \frac{\partial \mathcal{L}_0}{\partial (\partial_\mu \partial_\sigma \varphi)} + \dots$$

Then the gauge transformations are given by:

$$(3.250) \quad \delta_\epsilon \varphi^i(x) = \bar{R}_\alpha^i \epsilon^\alpha(x) + \bar{R}_\alpha^{i\mu} \partial_\mu \epsilon^\alpha + \cdots + E_\alpha^{i\mu_1 \cdots \mu_s} \partial_{\mu_1} \cdots \partial_{\mu_s} \epsilon^\alpha$$

and the rule

$$(3.251) \quad \delta_\epsilon \mathcal{L}_0 = \partial_\mu K_\epsilon^\mu \quad K_\epsilon^\mu(\varphi, \partial\varphi, \dots, \epsilon, \dots, \partial\epsilon)$$

It will be convenient to use some of the compact DeWitt notation here. Using this we can simplify:

$$(3.252) \quad \partial_\epsilon \varphi^i = R_\alpha^i \epsilon^\alpha \quad \delta_\epsilon \varphi^i(x) = \int d^D y R_\alpha^i(x, y) \epsilon^\alpha(y)$$

where

$$(3.253) \quad R_\alpha^i(x, y) = \bar{R}_\alpha^i \delta(x - y) - \bar{R}_\alpha^{i\mu} \partial_\mu \delta(x - y)$$

These symmetries give us the so-called local Noether identities. We know we have:

$$(3.254) \quad \delta S_0 = 0 = \frac{\delta S_0}{\delta \varphi^i} \delta_\epsilon \varphi^i = \frac{\delta S_0}{\delta \varphi^i} R_\alpha^i \epsilon^\alpha$$

and since this holds for any ϵ , this implies:

$$(3.255) \quad \frac{\delta S_0}{\delta \varphi^i} R_\alpha^i = 0$$

Note that if we don't want to use the compact notation this is written:

$$(3.256) \quad \frac{\partial \mathcal{L}_0}{\partial \varphi^i} \bar{R}_\alpha^i - \partial_\mu \left(\frac{\delta \mathcal{L}_0}{\delta \varphi^i} \bar{R}^{i\mu} \right) + \cdots = 0$$

THEOREM 3.3. *All such infinitesimal gauge transformations always form a Lie algebra.*

PROOF. This is the claim that these commutators:

$$(3.257) \quad [\delta_\eta, \delta_\epsilon] = f_{\eta\epsilon}{}^\chi \delta_\chi$$

satisfy the Jacobi identity. This is trivial from (3.255). \square

Note that we take $\bar{\mathcal{G}}$ to be the object generated by all such infinitesimal gauge transformations.

So what does this mean about open algebras? Well this seems to imply all such algebras are open. So what about theories with algebras which are not open? To understand this, we introduce the notion of a "trivial" gauge transformation.

Take a matrix of gauge transformations $\mu^{ij} = -\mu^{ji}$ where

$$(3.258) \quad \delta_\mu \varphi^i = \mu^{ij} \frac{\delta S_0}{\delta \varphi^j}$$

clearly these are indeed gauge transformations since

$$(3.259) \quad \delta_\mu S = \mu^{ij} \frac{\delta S_0}{\delta \varphi^i} \frac{\delta S_0}{\delta \varphi^j} \equiv 0$$

except these also don't imply any Ward identities. Their action on the fields is just proportional to the equations of motion. This means they don't constrain the action or imply anything relevant about the dynamics. Then we have the following theorem:

THEOREM 3.4. *The group \mathcal{N} generated by trivial gauge transformations, is a normal subgroup in $\bar{\mathcal{G}}$.*

PROOF. This can be shown explicitly for some transformation τ :

$$(3.260) \quad \delta_\tau \varphi^i = \tau^i \quad [\delta_\mu, \delta_\tau] \varphi^i = \left(\frac{\delta \tau^i}{\delta \varphi^k} \mu^{kj} - \frac{\delta \tau^j}{\delta \varphi^k} - t^k \frac{\delta \mu^{ij}}{\delta \varphi^k} \right) \frac{\delta S^0}{\delta \varphi^i}$$

□

So now we have two possibilities:

- (1) $\bar{\mathcal{G}} = \mathcal{N} \rtimes \mathcal{G}$ which means we can pick a particular embedding of $\mathcal{G} \hookrightarrow \bar{\mathcal{G}}$. Then \mathcal{G} is a Lie algebra since the commutators descend from $\bar{\mathcal{G}}$.
- (2) Alternatively $\bar{\mathcal{G}}$ is not a semidirect product, so we are stuck with $\bar{\mathcal{G}}$. This will be the case which leads to the so-called open symmetry algebra.

Proposition 3.5. *Not all Noether identities are independent.*

REMARK 3.3. What separates the physical gauge transformations from the trivial transformations are whether or not they give us independent Noether identities. This is because these constrain the structures of the correlation functions.

EXAMPLE 3.10. If we take $\delta_\epsilon \varphi^i = R_\alpha^i \epsilon^\alpha$ and define a different transformation:

$$(3.261) \quad \delta_\eta \varphi^i = R_\alpha^i M_A^B(\varphi) \eta^A$$

for an arbitrary matrix M , then trivially:

$$(3.262) \quad \frac{\delta S_0}{\delta \varphi^i} R_\alpha^i M_A^\alpha = 0$$

Define the so-called *generating set* G in $\bar{\mathcal{G}}^{3.3}$ to be any G which generates all of the Noether identities.

WARNING 3.3. Such a G is not necessarily even a subalgebra.

Proposition 3.6. *G is a generating set if any gauge transformation (that is any $\delta \varphi^i$ such that*

$$(3.263) \quad \delta \varphi^i \frac{\delta S_0}{\delta \varphi^i} = 0$$

is satisfied) can be written

$$(3.264) \quad \delta \varphi^i = \lambda^\alpha(\varphi) R_\alpha^i + M^{ij} \frac{\delta S_0}{\delta \varphi^j}$$

for $R_\alpha^i \in G$ and such that $M^{ij} = -M^{ji}$.

We will take this characterisation as an effective definition.

Proposition 3.7. *For any two elements of G ,*

$$(3.265) \quad \left[R_\alpha^i, R_\beta^j \right]$$

is a gauge transformation as well, so indeed this must be expressible in the form:

$$(3.266) \quad \delta \varphi^i = \lambda^\alpha(\varphi) R_\alpha^i + M^{ij} \frac{\delta S_0}{\delta \varphi^j}$$

^{3.3} If we were able to reduce to \mathcal{G} using the semidirect product, then we would define G within \mathcal{G} .

This means we can write:

$$(3.267) \quad R_\alpha^j \frac{\delta R_\beta^i}{\delta \varphi^j} - R_\beta^j \frac{\delta R_\alpha^i}{\delta \varphi^j} = C_{\alpha\beta}^\gamma(\varphi) R_\gamma^i + M_{\alpha\beta}^{ij}(\varphi) \frac{\delta S_0}{\delta \varphi^j}$$

So this is a general expression, but we can now consider some special cases.

- (1) Open vs. closed: If $M_{\alpha\beta}^{ij} \neq 0$, then G must be an open symmetric algebra.
- (2) Reducible vs. irreducible: If there exists some λ^α such that

$$(3.268) \quad \lambda^\alpha R_\alpha^i = N^{ij} \frac{\delta S_0}{\delta \varphi^j}$$

for $N^{ij} = -N^{ji}$ then G is called reducible. We can manually remove such cases to get a new G such that no such λ exist so G would be called irreducible.

EXAMPLE 3.11. Consider topological YM. Start with a gauge field A_μ^a . Then we have

$$(3.269) \quad \delta A_\mu^a = \delta_\mu^a(x) + D_\mu f^a(x)$$

which led us to introduce the ghost $\psi_\mu^a(x)$ and the chose $c^a(x)$. But we can just gauge this away by setting $\epsilon_\mu = -D_\mu f$. This is an example of a reducible G . To make this theory irreducible, we would have to introduce these ghost-for-ghosts $\varphi^a(x)$.

So now we can determine the smallest possible generating set G , and try to calculate commutation relations, to see what this means for the structure of G . In particular, we are interested to see if G is a proper Lie algebra. First we consider some examples:

EXAMPLE 3.12. Consider the abelian Chern-Simons gauge theory in $2+1$ dimensions. We can write the action:

$$(3.270) \quad S = \int A \wedge F = \int d^3x \epsilon^{\mu\nu\sigma} F_{\mu\nu} A_\sigma$$

this has some gauge symmetries, and we are interested in which such symmetries give independent Noether identities.

We claim this theory is invariant under two symmetries. It is certainly invariant under $U(1)$:

$$(3.271) \quad \delta A_\mu = \partial_\mu A(x)$$

and diffeomorphisms of our spacetime M :

$$(3.272) \quad \delta A_\mu = \xi^\rho \partial_\rho A_\mu + \partial_\mu \xi^\rho A_\rho$$

The difference between these two is a trivial gauge transformation. In particular we can show:

EXERCISE 7.1. Show that these two transformations only differ by a term which is proportional to $F^{\mu\nu}$.

This means they differ by something which gives us a trivial Ward identity.

EXAMPLE 3.13. Consider point-like particle quantum mechanics:

$$(3.273) \quad S_{QM} = \int dt (p\dot{q} = H(p, q))$$

which has a large amount of gauge symmetry for any Hamiltonian H . In particular, this is invariant under:

$$(3.274) \quad \delta q = \epsilon(t) \left(\dot{q} - \frac{\partial H}{\partial p} \right) \quad \delta p = \epsilon(t) \left(\dot{p} + \frac{\partial H}{\partial q} \right)$$

which are in $\overline{\mathcal{G}}$. In fact we can show that the algebra of the $\epsilon(t)$ is isomorphic to $\text{Diff}(\mathbb{R})$. These are in fact also in our normal subgroup \mathcal{N} , so the statement here should really be that there is no gauge symmetry, since these have no implication for the dynamics of the theory. That is, \mathcal{G} is trivial.

7.3. Cohomological approach. So we have seen that we should not be thinking of $\overline{\mathcal{G}}$ as the gauge group. We should first try for \mathcal{G} , and if not we should reduce to this minimal G . The cost we might pay would be that $[g, g']$ doesn't close to a Lie algebra. We consider only the irreducible case. Then we can define the *stationary surface* to be the surface Σ where the equations of motion:

$$(3.275) \quad \frac{\delta S_0}{\delta \varphi^i} = 0$$

are satisfied. We have the following theorem:

THEOREM 3.8. *Generators X_i of G*

$$(3.276) \quad [X_i, X_j] \propto X_k$$

form an integrable distribution on this stationary surface Σ in the sense that they generate gauge orbits.

We want to determine what it means for something to be an observable in this context. As we have often done, we now briefly return to the Hamiltonian formalism to make sense of things here. We know that in Hamiltonian QM, the observables are C^∞ functions on phase space. What about for some complicated quantum field theory? In particular, for some theory with gauge invariance, what is the appropriate phase space? Well in some sense we want to integrate over all possible histories, so we want to consider initial conditions for the equations of motion, which are equivalent to solutions of the equations of motion. We take the covariant phase space to be the space Σ of classical solutions. From the previous theorem, we should be identifying observables with functions on the space of orbits in Σ which is clearly Σ/G . So we have the space of all fields Φ on Σ , then the orbits of G on Σ , which gives a projection to Σ/G .

But how do we identify the collection of smooth functions $C^\infty(\Sigma/G)$? To do this we define a charge s which plays the role of the BRST charge Q . That is, we want $s^2 = 0$ to use the tools of cohomology. We introduce a grading by the ghost number. In particular, we take the ghost number of s to be 1, and we take

$$(3.277) \quad H^0(s) = C^\infty(\Sigma/G)$$

We will also find the novelty that

$$(3.278) \quad sA = (A, S)$$

for any observable (or more generally any function) A where S starts as the classical action of the theory and gets quantum corrections.

So now we desire to see this from first principles. First we localize to $C^\infty(\Sigma)$, and second we incorporate G . The first step is resolved using the Koszul-Tate resolution. For any operator δ , we have

$$(3.279) \quad H_0(\delta) = \frac{\ker \delta}{\text{im } \delta} = C^\infty(\Sigma)$$

Functions from this space which vanish on the surface Σ form an ideal which we call \mathcal{I} . Then we have

$$(3.280) \quad C^\infty(\Sigma) = C^\infty(\Phi) / \mathcal{I}$$

So if we want to identify a δ which gives us this space, we need to construct δ such that:

$$(3.281) \quad \ker \delta = C^\infty(\Phi) \quad \text{im } \delta = \mathcal{I}$$

so define $\delta\varphi^i = 0$ for the first requirement, and for the second, for each φ^i we introduce an anti-field φ_i^* , where we take:

$$(3.282) \quad \delta\varphi_i^* = -\frac{\delta S_0}{\delta\varphi^i}$$

which gives us such a δ . Clearly since δ has ghost number 1, $\text{gh}\#(\varphi_i^*) = -1$. We also define an antighost number which assigns $\text{anti-gh}\#(\varphi_i^*) = 1$. Now in the absence of gauge symmetry, $H_p(\delta) = 0$. However if we have any nontrivial gauge symmetry, and therefore a Noether identity, we have a nontrivial equation:

$$(3.283) \quad \frac{\delta S_0}{\delta\varphi^i} R_\alpha^i = 0$$

which will ruin this identically zero homology. The RHS in (3.282) is degenerate in this case. Now by this identity, we can in fact write:

$$(3.284) \quad \delta(R_\alpha^i \varphi_i^*) = 0$$

so this is a nontrivial cohomology element. So now introduce φ_α^* such that

$$(3.285) \quad \delta\varphi_\alpha^* = R_\alpha^i \varphi_i^*$$

We introduce one of these for every element of G . Since $\text{gh}\#(\delta) = 1$, $\text{anti-gh}\#(\varphi_\alpha^*) = 2$. So if we are in the irreducible case, then there will be a potential obstacle at antighost number 2. So we have to introduce some objects with antighost number 3 to kill this nontrivial cohomology. We do not pursue this further.

Consider only the space Σ . Now we treat this as having a fibration as given by the orbits of G . So we project each orbit to this space Σ/G . So the punchline here is that this reduction $C^\infty(\Sigma/G)$ is just the old BRST procedure. To see this we define vertical vectors to be the vectors on Σ tangent to the orbits \mathcal{O} . We also define vertical p -forms to be p -forms along the orbits. Then we want to define an exterior derivative d on these forms such that it takes vertical p -forms to vertical $p+1$ -forms. For $p=0$ define

$$(3.286) \quad dF(X) = \frac{\partial}{\partial X} F$$

If we have α some vertical 1-form we define:

$$(3.287) \quad d\alpha(X, Y) = -\mathcal{L}\alpha(x) - \mathcal{L}_x\alpha(y) + \alpha([X, Y])$$

where \mathcal{L} is the Lie derivative. Then we claim $d^2 = 0$, and that d plays the role of the BRST charge. Now pick a basis of vertical vectors:

$$(3.288) \quad X_\alpha F = \frac{\delta F}{\delta\varphi^i} R_\alpha^i$$

which satisfy commutation relations:

$$(3.289) \quad [X_\alpha, X_\beta] = C_{\alpha\beta}^\gamma(\varphi) X_\gamma$$

on Σ . So now introduce fields $\{c^\alpha\}$ called ghosts to be dual to the $\{X_\alpha\}$. These are indeed the same ghosts from the typical BRST story. Now define

$$(3.290) \quad dF = (X_\alpha F) c^\alpha \quad dc^\alpha = \frac{1}{2} c_{\alpha\beta}^\gamma c^\beta c^\gamma$$

So we want the BRST charge to satisfy both of these. That is, we want some differential which satisfied the notable qualities of both δ and d . We might guess it is $\delta + d$, but $\delta d + d\delta \neq 0$, so we need to add more terms. This process is called Homological perturbation theory. But there is actually a shortcut to showing this order by order. It turns out the action of s is just given as the antibracket which is given by some function A and acts on S :

$$(3.291) \quad sA = (A, S)$$

So we have fields φ^i and φ_j^* and we define

$$(3.292) \quad (\varphi^i, \varphi_j^*) = \delta_j^i \quad (c^\beta, \varphi_\alpha^*) = \delta_\alpha^\beta \quad \dots$$

which would continue if we had ghosts for ghosts etc. Now we insist on the following three properties:

- (1) $\text{gh}\#((A, B)) = \text{gh}\#(A) + \text{gh}\#(B) + 1$
- (2) $\epsilon((A, B)) = \epsilon(A) + \epsilon(B) + 1$
- (3) $(A, B) = -(-1)^{(\epsilon_A+1)(\epsilon_B+1)}(B, A)$
- (4) Jacobi: $(-1)^{(\epsilon_A+1)(\epsilon_B+1)}(A, (B, C)) + \text{cyclic permutations} = 0$

In addition we have $\text{gh}\#(S) = 0$ and $\epsilon(S) = 0$ along with what is called the master equation:

$$(3.293) \quad (S, S) = 0$$

So finding a solution for the master equation amounts to finding such an operator s . We can also write S as a sum of various objects with various ghost numbers:

$$(3.294) \quad S = \sum_{m \geq 0} S^{(m)}$$

Then S has the following properties:

- (1) $S^{(0)} = S$
- (2) $(\psi_i^*, S(\varphi_i)) = \delta\varphi_i^*$
- (3) $S^{(1)} = \varphi_i^* R_\alpha^i c^\alpha$

7.4. Quantization. Now we can finally perform quantization of this theory. The first step is to introduce the antighosts b_A, B_A . But now we have to introduce antifields b_A^* and B_A^* in anticipation of the gauge fixing terms. Then we simplify notation by introducing a collective name

$$(3.295) \quad \varphi^{\bar{A}} = (\varphi^i, c^\alpha, b_A, B_A, \dots) \quad \varphi_{\bar{A}}^* = (\varphi^{i*}, c^{\alpha*}, b_A^*, B_A^*, \dots)$$

Then define

$$(3.296) \quad \mathcal{Z}^{\bar{a}} = (\varphi^{\bar{A}}, \varphi_{\bar{B}}^*)$$

and the antibracket on these to be:

$$(3.297) \quad (A, B) = \frac{\delta^r A}{\delta \mathcal{Z}^{\bar{a}}} \zeta^{\bar{a}\bar{b}} \frac{\delta^l B}{\delta \mathcal{Z}^{\bar{b}}}$$

where

$$(3.298) \quad \mathcal{Z}^{\bar{a}\bar{b}} = \begin{pmatrix} 0 & \delta_{\bar{B}}^{\bar{A}} \\ -\delta_{\bar{B}}^{\bar{A}} & 0 \end{pmatrix}$$

Now we have N gauge symmetries which still exist, so we set

$$(3.299) \quad \Omega^{\bar{A}}(\mathcal{Z}^a) = 0$$

while maintaining $s^2 = 0$. Then it turns out that it is equivalent to satisfy the involution:

$$(3.300) \quad \left(\Omega^{\tilde{A}}, \Omega^{\tilde{B}} \right) = 0$$

So if these satisfy this, then they all follow from one gauge fixing fermion $\Phi(\varphi^A)$ and we have expressions of the form:

$$(3.301) \quad \varphi_A^* = \frac{\delta \Psi}{\delta \varphi^A}$$

So the procedure to get back to our familiar BRST procedure is to write down the action for the theory, and substitute this gauge fixing condition.

We will be returning to this formalism to discuss its relationship with anomalies, but first we offer an overview of large N theories and the AdS / CFT correspondence.

Large N (AdS / CFT correspondence)

1. Historical context

We now consider a non-perturbative technique known as taking the “large N limit.” We aim to give some motivation and overview as to why this is called this and provide some historical context. The parallel we should really have in mind here is bosonization. Recall that in our treatment of this we first considered the Thirring model and we discovered a duality between this and the sine-Gordon theory.

The first real application of this formalism came in H.E. Stanley’s publication [22] in 1968. In particular, he proposed to start with a spin system Hamiltonian:

$$(4.1) \quad H = -J \sum_{\langle ij \rangle} \vec{s}_i \vec{s}_j$$

where \vec{s}_i is an $O(n)$ vector and we are summing over nearest neighbors on some lattice. These are surprisingly rich statistical theories. For $n = 0$ we have self avoiding walks on the lattice, $n = 1$ is the so-called Ising model, $n = 2$ gives the XY model, and $n = 3$ is the Heisenberg model. Now Stanley suggests that we take a $1/n$ expansion. In order for this to be a legitimate expansion, we of course have to take $n \rightarrow \infty$. As a part of this process, we have to decide what we keep fixed as we take this limit. For example we might allow J to scale.

The large N approach was brought to high energy theory by G. ’t Hooft in 1974 [13]. Recall the action from QCD:

$$(4.2) \quad S_{\text{QCD}} = \frac{1}{2g_{\text{ym}}^4} \int \text{tr}(F_{\mu\nu} F^{\mu\nu}) + \int \bar{\Psi} \not{D} \Psi + m \bar{\Psi} \Psi$$

but we have seen that this theory exhibits dimensional transmutation. That is g_{ym} turns itself into a function of the scale, which means all we really have is a characteristic energy scale of the theory. So we don’t have a free renormalized coupling, but rather we end up with Λ_{QCD} which is the scale where $g_{\text{ym}}(\Lambda) \sim 1$. But we can’t expand in terms of Λ , so we cannot use perturbative techniques in this way.

But ’t Hooft suggested we can consider a different parameter. For this $F_{\mu\nu}$ we have the gauge group $G = \text{SU}(N)$ Now we can use $1/N$ as an expansion parameter as Stanley did for statistical mechanics. Again we have to take a large N limit. It turns out this is even more practical than bosonization because it is easier to generalize to arbitrary quantum field theories.

2. 1 + 1 dimensions

Recall the Thirring model which we covered in our treatment of bosonization. On some flat spacetime, we have some self interacting fermions:

$$(4.3) \quad S_T = \int_{\mathbb{R}^2} d^2x \left\{ \bar{\psi} i \not{\partial} \psi - g_T (\bar{\psi} \gamma_\mu \psi)^2 \right\}$$

We consider a similar model called the Gross-Neveu model:

$$(4.4) \quad S_{GN} = \int_{\mathbb{R}^2} d^2x \left\{ \bar{\psi} i \not{\partial} \psi - \frac{g}{2} (\bar{\psi} \psi)^2 \right\}$$

where ψ^a running from $a \in \{1, \dots, N\}$ are Dirac spinors. This has a global $SU(n)$ flavor symmetry, which extends to $SO(2n)$, but we are only concerned with keeping track of the $SU(n)$ symmetry. Note that in this case we of course have $[g] = [g_T] = 0$. Classically this theory has a discrete $\mathbb{Z}/2\mathbb{Z}$ chiral symmetry. This action is given by the chirality matrix γ_3 which acts as:

$$(4.5) \quad \psi \rightarrow \gamma_3 \psi \qquad \bar{\psi} \rightarrow -\bar{\psi} \gamma_3$$

This theory also has no classical mass. But just as we saw before, we might meet some non-perturbative techniques to get that g is asymptotically free. Then this may lead to the so-called “running” of the coupling constant which give us a non-perturbative mass scale.

For $N = \infty$ and no $1/N$ corrections we have the following Feynman rules:

$$(4.6) \quad a \longrightarrow b \sim \delta_{ab}$$

$$(4.7) \quad g \sim \begin{array}{c} \begin{array}{ccc} a & & a \\ & \searrow & / \\ & \times & \\ & / & \searrow \\ b & & b \end{array} & \equiv & \begin{array}{ccc} a & & a \\ & \searrow & / \\ & \vdots & \\ & / & \searrow \\ b & & b \end{array} \end{array}$$

Now we consider the scattering of $a, \bar{a} \rightarrow b, \bar{b}$.

$$(4.8) \quad \begin{array}{c} \uparrow \\ t \end{array} \begin{array}{ccc} & b & \\ & \searrow & / \\ & \times & \\ & / & \searrow \\ a & & a \end{array}$$

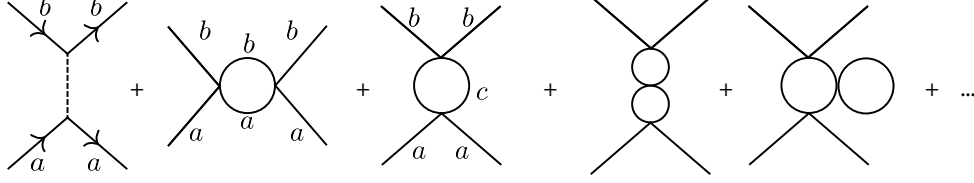
Now there is no naive way to take $N \rightarrow \infty$. We have to rescale the parameters in the theory as we take this limit. Typically there is a unique way to do this such that the theory is nontrivial and finite. We fix the so-called 't Hooft coupling $\lambda^2 = gN$.

Now we have

$$(4.9) \quad S = \int d^2x \left\{ \bar{\psi}^a i \not{\partial} \psi^a - \frac{\lambda^2}{2N} (\bar{\psi} \psi)^2 \right\}$$

It might be interesting to integrate in some auxiliary field $\sigma(x)$. We can then rewrite this action as:

$$(4.10) \quad S = \int d^2x \left\{ \bar{\psi} i \not{\partial} \psi - \sigma \bar{\psi} \psi + \frac{N}{2\lambda^2} \sigma^2 \right\}$$

FIGURE 1. Loop diagrams for the $a, \bar{a} \rightarrow b, \bar{b}$ scattering:

since we can integrate out σ to get the original action. At the classical level, the EOM for σ is just:

$$(4.11) \quad \sigma = \frac{\lambda}{N} \bar{\psi} \psi$$

We get an honest propagator for this field:

$$(4.12) \quad \sigma \sim \dots \rightarrow \dots$$

but now we can integrate $\psi, \bar{\psi}$ out with a Grassmanian Gaussian. This is perturbatively equivalent to summing an infinite sequence of diagrams to obtain an effective potential for ψ . This can be seen in fig. 1. Notice that the orders of the terms in fig. 1 are given by:

$$(4.13) \quad \mathcal{O}(g) = \mathcal{O}(\lambda^2/N) \quad \mathcal{O}(g^2) = \mathcal{O}(\lambda^4/N^2)$$

$$(4.14) \quad \mathcal{O}(g^2 N) = \mathcal{O}(\lambda^4/N) \quad \mathcal{O}(g^3 N^2) = \mathcal{O}(\lambda^6/N)$$

$$(4.15) \quad \mathcal{O}(g^3 N^2) = \mathcal{O}(\lambda^6/N)$$

Explicitly we calculate:

$$(4.16) \quad -iV(\sigma_0) = -i \frac{N}{2\lambda^2} \sigma_0^2 - N \sum_{r=1}^{\infty} \frac{1}{2r} \text{Tr} \int \frac{d^2 p}{(2\pi)^2} \left(-\frac{\not{p} \sigma_0}{p^2} \right)^{2r}$$

$$(4.17) \quad V = \frac{N}{2\lambda^2} \sigma^2 - N \int \frac{d^2 p}{(2\pi)^2} \log \left(1 + \frac{\sigma_0}{p^2} \right)$$

Now doing dimensional regularization, \overline{MS} scheme, we get

$$(4.18) \quad V = N \left\{ \frac{\sigma^2}{2\lambda^2} + \frac{\sigma_0^2}{4\pi} \left(\log \frac{\sigma_0^2}{\mu^2} - 1 \right) \right\}$$

This also satisfies a renormalization group equation:

$$(4.19) \quad \left\{ \mu \frac{\partial}{\partial \mu} + \beta(\lambda) \frac{\partial}{\partial \lambda} - \gamma_\sigma(\lambda) \sigma \frac{\partial}{\partial \sigma} \right\} V(\sigma) = 0$$

where γ_σ is the anomalous dimension of σ . Notice this is satisfied for $\gamma_\sigma(\lambda) = 0$ and $\beta(\lambda) = -\lambda^3/(2\pi)$. Now the novelty is that this result is exact to all λ orders as $N \rightarrow \infty$ so we have found the exact β function. We can also notice from the sign of β that the theory is asymptotically free. Finally we get that there is indeed dimensional transmutation

present in this theory. So even though the theory originally had no scale, through the RG process, we develop a dependence on an effective scale. One way of showing this is to just fish for the minima of the effective potential. There is a naive one at $\sigma = 0$, and then at $\langle\sigma\rangle = \pm\mu e^{-\pi/\lambda^2}$. And since this is nonzero, this is the analogue of Λ_{QCD} from before.

REMARK 4.1. Any time we have $e^{-\#/G^2}$ for some coupling G , this expression around $G = 0$ has the trivial expansion 0. This is a hallmark of a nonperturbative effect, since it cannot be seen to any order in perturbation theory.

Now returning to our effective potential, we can substitute $\langle\sigma\rangle$ for σ and we get chiral symmetry breaking.

Recall that we just made a guess for this form of σ . This was really only easy to guess here. Treating something more complicated such as QCD makes this much more difficult.

Also notice, that if we write the theory in the language of σ , then we can treat $\hbar \sim 1/N$. This is justified by the original path integral when we have integrating out ψ :

$$(4.20) \quad \int \mathcal{D}\bar{\psi}\mathcal{D}\psi\mathcal{D}\sigma e^{-S(\psi,\bar{\psi},\sigma)} \rightarrow \int \mathcal{D}\sigma e^{-S_{\text{eff}}(\sigma)}$$

where

$$(4.21) \quad S_{\text{eff}}(\sigma) = N\tilde{S}_{\text{eff}}(\sigma, \lambda)$$

If we didn't rewrite the theory in the language of σ so we just keep the original ψ variables, we might wonder if we would have been able to show the same thing? Well we can take $\Psi = \psi/\sqrt{N}$ and get

$$(4.22) \quad S_{\text{GN}} = N \int \Psi^a \not{\partial} \Psi - \lambda^2 \left(\bar{\Psi}^a \Psi^a \right)^2$$

but there is this implicit dependence on N within the integral given by the index a . Since σ was an $\text{SU}(n)$ singlet, we didn't have the dependence on N within the integral before.

3. Riemannian surfaces

Before we apply this methodology to a general gauge theory, we develop some formalism. Consider all two-dimensional orientable surfaces. We desire to classify such surfaces. As it turns out, we can classify these surfaces by their genus. This is effectively just the number of handles. We can see this in fig. 2.

We can also classify compact orientable surfaces with the Euler number. For a surface Σ of genus g , this is given by $\chi(\Sigma) = 2 - 2g$. Without appealing to the genus, we can write this as:

$$(4.23) \quad \chi(\Sigma) = \sum (-1)^k \dim H_{\text{de Rham}}^k(\Sigma)$$

Recall the following De Rham cohomologies:

$$(4.24) \quad H^0 = \mathbb{R} \qquad H^1 = \mathbb{R}^2 \qquad H^2 = \mathbb{R}$$

As an example we can see that $\chi(T^2) = 0$. We can also calculate this by integrating the curvature:

$$(4.25) \quad \chi(\Sigma) = \frac{1}{4\pi} \int d^2\sigma \sqrt{g} R$$

As a final method for calculating the Euler characteristic, we can consider the simplicial decomposition of a surface. This just means we compose the surface into its irreducible elements. So we preserve the space topologically, but decompose it into vertices, edges,

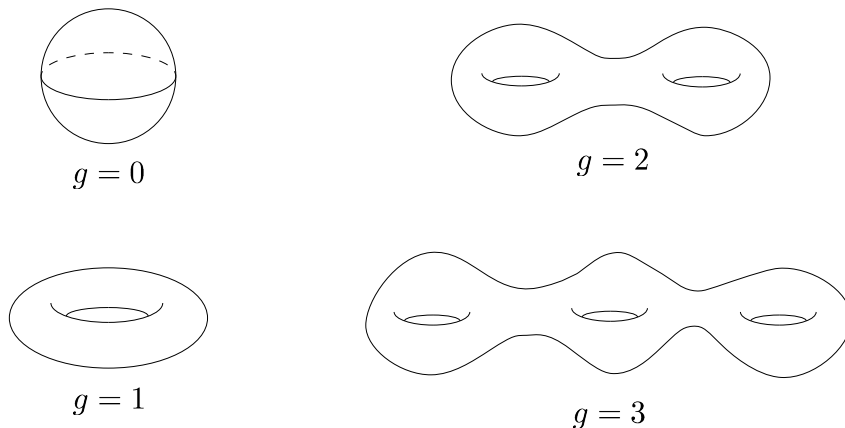


FIGURE 2. Illustrations of the genus of a Riemann surface.

and plaquettes. For example, the sphere can be topologically deformed to a tetrahedron, which provides the explicit simplicial decomposition where we have 4 vertices, 6 edges, and 4 surfaces. Then the Euler characteristic is given by:

$$(4.26) \quad \chi(\Sigma) = \#V - \#E + \#L$$

where L gives the number of closed loops, or plaquettes. We can check that for the sphere this agrees with the value of $\chi(S^2) = 2$ that we have seen via other methods.

DEFINITION 4.1. A planar graph is a graph which can be embedded into a surface of genus 0 without crossing edges.

In other words, a planar graph is just a graph that can be drawn on a sphere or plane. As motivated by the definition of a planar graph and the simplicial decomposition, if a finite, connected, planar graph is drawn in the plane without any edge intersections, then the Euler characteristic of a planar graph (which is just $v - e + f$) must be 2.

4. Generic treatment

We now treat a generic QFT with some fields M , which are matrices of some Lie group such as $U(N)$, $SU(N)$, $SO(N)$, $Sp(n)$ etcetera. We take $U(N)$ for simplicity. We label this matrix as M_j^i , and assume it is Hermitian. We have the $U(1)$ action (gauge or global, this argument doesn't care) given by:

$$(4.27) \quad U(1) : M \rightarrow U M U^\dagger$$

We want to treat any theory with a Lagrangian which we can write in terms of M as:

$$(4.28) \quad S = \frac{1}{g_c^2} \int \text{Tr} \left\{ (\partial M)^2 + \underbrace{M^2 + M^3 + \dots}_{V(M)} \right\}$$

We write g_c for the coupling constant for reasons which will become obvious below. Now we make these large N -limit considerations. We will see that every such field theory under this

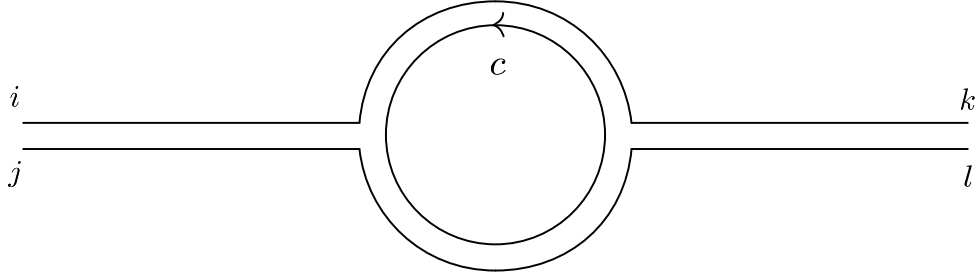


FIGURE 3. Example of a one-loop ribbon diagram.

large- N expansion is a string theory. In particular, we eventually define $1/N$ as the string coupling constant.

Now we write Feynman diagrams, and determine which loops can be kept or discarded in this limit. Since M now has two indices, each propagator looks like a ribbon here, and the Feynman diagrams become ribbon diagrams.

$$(4.29) \quad \begin{array}{c} i \longleftarrow l \\ j \longrightarrow k \end{array} \sim \langle M_j^i M_l^k \rangle$$

REMARK 4.2. If we were using $\text{SO}(n)$ rather than $\text{SU}(n)$ we would have that these are unoriented, so these would just be double lines.

We can think of these as thickened lines. Now what happens as $N \rightarrow \infty$? Take P to be the number of propagators. This goes as g_c^2 . Each vertex will go as $1/g_c^2$. Each loop L goes as N . See fig. 3 for an example of such a loop. Now for each diagram, just from the combinatorics we get:

$$(4.30) \quad (g_c^2)^{\#P-\#V} N^{\#L} = (g_c^2 N)^{\#P-\#V} N^{\#V-\#P+\#L}$$

This leads us to anticipate a t' Hooft coupling $\lambda = g_c^2 N$. We now claim that:

$$(4.31) \quad N^{\#V-\#P+\#L} = \left(\frac{1}{N}\right)^{2g-2}$$

for some $g \in \mathbb{N}$ which will eventually be the genus of some surface.

For a general ribbon diagram, we can consider it as a graph. As such, we can consider whether or not it is planar. To see this, we fill in all of the ‘‘closed’’ loops with a disk, and see if we cover the full plane. Alternatively we can check if any edges cross. So in fig. 3 we fill in the inner circle and the portions above and below, so this is clearly a planar graph. In fig. 4 on the left, we see another example of a planar graph. Also notice that as mentioned in the previous section, this graph does indeed have Euler characteristic $2 = 2 - 3 + 3$. However in fig. 4 on the right, this is not a planar graph. Note that it has Euler characteristic $4 - 6 + 2 = 0 \neq 2$. It can however be drawn on the torus T^2 . Then we have the following statement:

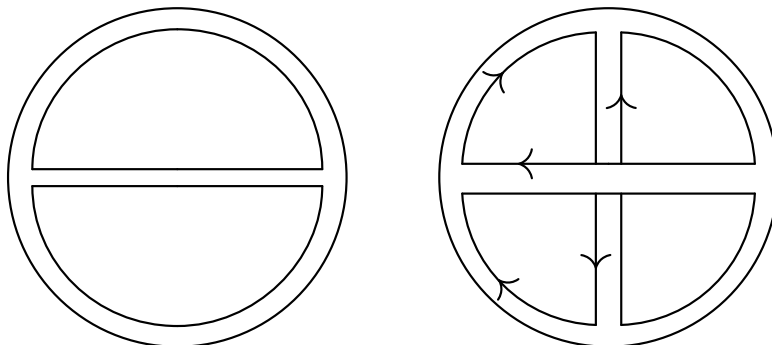


FIGURE 4. More examples of ribbon diagrams.

THEOREM 4.1. *Each such ribbon diagram is associated uniquely with a Riemann surface of some genus g .*

Now take

$$(4.32) \quad \mathcal{Z} = \sum_{\text{diag.}} \left(\frac{1}{N} \right)^{2g-2} = \sum_{g=0}^{\infty} \left(\frac{1}{N} \right)^{2g-2} \sum_{\text{diag. of genus } g} f(\lambda)$$

$$(4.33) \quad = \sum_{g=0}^{\infty} g_{\text{string}}^{2g-2} F_g(\lambda)$$

where as we mentioned in the beginning, we have taken $1/N$ as our string coupling g_{string} . A string theory can be pragmatically defined to be a theory which rather than an expansion into graphs with 1-dimensional propagator lines, has an expansion of Feynman graphs with the geometry of 2-dimensional surfaces.

5. AdS space

So we have this correspondence between large N theories of matrices and string theories. But we don't have a good prescription of calculating these F_g coefficients. Indeed, very few "string theories" are known. So which large N theories of M^i_j can be described by known string theories? This is essentially what led Maldacena to the AdS/CFT correspondence in 1997 [17].

In particular we seek quantum field theories with RG flow, and Poincaré symmetry. We follow our procedure from the past, and restrict our attention to the fixed points of the RG. It is easier to first classify these fixed points and then classify the flow in the RG. These fixed points will have an extension of the Poincaré symmetry to conformal symmetry. For example, $\text{SO}(3, 1) \times \mathbb{R}^4$ extends to the conformal symmetry $\text{SO}(4, 2)$. Then perhaps, we can use this symmetry to more uniquely identify a candidate for the dual string theory.

It is beautiful to have this idea, but how many such points do we know besides free field theories? One possible trick to get further, is to add SUSY. In particular, $N = 4$ SUSY in $3 + 1$ dimensions was shown to be an exact CFT by Stanley Mandelstam. So we have the fixed point of RG given by c^μ for $\mu = 0, \dots, 3$, and we have this enhanced global $\text{SO}(4, 2)$.

So we have the equation

$$(4.34) \quad \mathcal{Z}(\lambda, N) = \sum_{g=0}^{\infty} g_{\text{string}}^{2g-2} F_g(\lambda)$$

If we can define these independently of one another, then they share the same Hilbert space. We know $\text{SO}(4, 2)$ is a conformal symmetry of $\mathbb{R}^{3,1}$, but how do we realize this $\text{SO}(4, 2)$ on the string theory side? We call the Riemann surface the string worldsheet, and postulate:

$$(4.35) \quad F_g(\lambda) = \int \mathcal{D}(\text{fields}) e^{iS_{WS}}$$

where we have

$$(4.36) \quad S_{WS} = \int d^2\sigma \mathcal{L}(f, \lambda, \dots)$$

So say we have the worldsheet surface Σ_g with coordinates σ . The simplest way to associate some fields with this string theory, is to consider this spacetime \mathcal{M} with coordinates x^μ , and consider maps from this Σ_g to \mathcal{M} . Then these are our candidate fields.

The simplest worldsheet Lagrangian^{4.1} is:

$$(4.37) \quad S_{WS} \sim \int d^2\sigma \partial_\alpha x^\mu \partial_\alpha x^\nu \eta_{\mu\nu}$$

If we stare at this, we can see that it is indeed $\text{SO}(3, 1)$ invariant, but is not $\text{SO}(4, 2)$ invariant. This seems to be very bad. As it turns out, the way to refine this Lagrangian, is to realize that the reason for this is that the target space \mathcal{M} only has Lorentz invariance. So if we want the $\text{SO}(4, 2)$, let's search for a target space $\hat{\mathcal{M}}$ which has $\text{SO}(4, 2)$ as an isometry group. This is easy enough to accomplish, since \mathcal{M} with coordinates c^μ already realizes our $\text{SO}(3, 1)$. If we want to extend this Poincaré symmetry into something bigger, we have to add dimensions.

The simplest way to do this is by brute force. Just introduce a new dimension u . There will be a metric on this bigger spacetime given by:

$$(4.38) \quad ds^2 = w^2(u) dx^\mu dx^\nu \eta_{\mu\nu} + w(u)^2 du^2$$

for some w . The first term has this Minkowski term up to a multiple, so it still gives us the $\text{SO}(3, 1)$. Now we insist that this metric has $\text{SO}(4, 2)$ isometries. So add conformal rescalings of \mathcal{M} : $x^\mu \rightarrow \lambda x^\mu$. This forces us to also rescale $u \rightarrow \lambda u$, which restricts our w . Namely, $w^2 = 1/u^2$ will produce the desired invariant metric. So our prediction is that the string theory which is dual to this CFT should live on a spacetime in 5-dimensions, with metric:

$$(4.39) \quad \frac{L^2}{u^2} (dx_\mu dx^\mu + du^2)$$

where L is called the radius.^{4.2} If this is set to 1, the metric is just:

$$(4.40) \quad ds^2 = \frac{1}{u^2} (dx^\mu dx_\mu + du^2)$$

This is a famous spacetime, called the anti-de Sitter (AdS) space in 4 + 1 dimensions with the Poincaré patch coordinates. We write this space as AdS_5 . We call this u coordinate the radial coordinate. We will see that the radial evolution is effectively the same as the RG

^{4.1} This is a cartoon version of what is called a ‘‘Polyakov action.’’

^{4.2} We don't use R because we will be discussing gravitational effects, and we don't want to confuse this with the Ricci scalar.

flow. The boundary $u = 0$ will correspond to the UV limit of a CFT. The IR limit is then just deep in the bulk of the geometry. In other words, UV in CFT corresponds to the IR in gravity.

So returning to our $N = 4$ SYM in $3 + 1$, we have 16 Q s in our full super-algebra, and 16 S s, meaning we have 32 super-symmetries. Now we get out $SO(4, 2)$, but we have more. Recall the Q_α^I all have R -symmetry index I , and each has 4. So in fact, we have the bigger symmetry:

$$(4.41) \quad SO(4, 2) \otimes SU(4)$$

As before, we still have this AdS space since we have the $SO(4, 2)$ symmetry, but now we have to introduce another manifold. Recall the accidental isomorphism $SU(4) \cong SO(6)$. So instead of realizing $SU(4)$, we can realize $SO(6)$, which acts as rotations on a 6 dimensional space. In particular, we predict that the string theory dual for this theory has to be a SUSY theory whose ground state which looks like AdS_5 , has 32 SUSY, and lives on $AdS_5 \times S^5$.

6. Establishing the dictionary

6.1. Correlation functions. Without even introducing SUSY, we can evaluate the gravity partition function as a function of the boundary values of the fields. We consider the bulk partition function:

$$(4.42) \quad \mathcal{Z}_{bulk} [\varphi(\vec{x}, u)|_{u=0} = \varphi_0(\vec{x})] = \left\langle \exp \left(\int d^4x \varphi_0(\vec{x}) \mathcal{O}(\vec{x}) \right) \right\rangle_{\text{CFT}}$$

where we have specified the boundary conditions for the fields on the boundary of AdS.

Now we will see that the dual string theory will always include gravitational effects. This means if we restrict our attention to theories of large N matrices with a local energy momentum tensor $T^{\mu\nu}$, this implies there is a $g_{\mu\nu}$ excitation on the string side. By this, we mean a propagating, massless, spin-2 particle.

EXAMPLE 4.1. In particular, we consider the example of a scalar field φ of mass m in bulk:

$$(4.43) \quad S = \frac{1}{2} \int d^4x \, du \, \sqrt{-G} \left(G^{MN} \partial_M \varphi \partial_N \varphi - \frac{1}{2} m^2 \varphi^2 \right)$$

For the moment we consider solutions near the boundary of AdS, which is given by $u \approx 0$. In particular, we look for solutions of the form $\varphi \sim u^\alpha$, where

$$(4.44) \quad \alpha(\alpha - 4) - m^2 L^2 = 0$$

then we can plug this back in, which restricts α to be:

$$(4.45) \quad \alpha_\pm = 2 \pm \sqrt{4 + m^2 L^2}$$

Now we have that the α_- solution dominates near $u \rightarrow 0$, and the α_+ solution always decays when $u \rightarrow 0$. This gives us the Breitenlohner-Freedman bound:

$$(4.46) \quad m^2 L^2 \geq -4$$

For φ , the \mathcal{O} has conformal dimension Δ , where

$$(4.47) \quad \Delta = 2 + \sqrt{4 + m^2 L^2}$$

6.2. Holographic renormalization. We have seen this expression:

$$(4.48) \quad \mathcal{Z}_{bulk} [\varphi(\vec{x}, u)|_{u=0} = \varphi_0(\vec{x})] = \left\langle \exp \left(\int d^4x \varphi_0(\vec{x}) \mathcal{O}(\vec{x}) \right) \right\rangle_{\text{CFT}}$$

but we have UV divergences on the RHS, meaning we need some regularization and renormalization. Using the UV IR correspondence, this means we have an IR divergence on the LHS. To treat this, we take $u = \epsilon$ for some small ϵ and take this to zero, to see how this blows up. The resulting procedure should reproduce the usual renormalization procedure on the CFT side of the correspondence.

So evaluate $\varphi(\vec{x}, u)$ at $u = \epsilon$, which yields:

$$(4.49) \quad \varphi(\vec{x}, u)|_{u=\epsilon} = \epsilon^{\alpha_-} \varphi_0^R(\vec{x})$$

What is the scaling (engineering) dimension of φ ? Well we have the rescalings: $\vec{x} \rightarrow \lambda \vec{x}$, $u \rightarrow \lambda u$, and $\varphi \rightarrow \varphi$. Therefore $[\varphi^R] = \alpha_-$. But then we have that $[\mathcal{O}] = 4 - \alpha_- = \alpha_+$, which means

$$(4.50) \quad \Delta = 2 + \sqrt{4 + m^2 L^2}$$

as we have already seen! We have only shown this for the scalar field, but this turns out to be true in general. Recall we have $[T^{\mu\nu}] = 4$ which is not getting renormalized. This implies there is a massless source field $J_{\mu\nu}$ in the bulk. We write this $g_{\mu\nu} \subset G_{\mu\nu}$. This is a massless spin-2 field, which must be embedded in a theory local spacetime with gauge invariance. In other words, this must be a graviton.

Now let's take a limit where the LHS is approximated well by classical gravity. In this limit, the effective action $\log \mathcal{Z}_{bulk} = W_{\text{eff}}$, will be

$$(4.51) \quad W_{\text{eff}} = -S_{\text{on-shell}} [\varphi_0^R]_{\text{gravity}}$$

We only keep track of the dependence of the space time metric $g_{\mu\nu}$, which we have just argued for the existence of. So $W_{\text{eff}} = -S_{\text{on-shell}} [g_{\mu\nu}]_{\text{gravity}}$ and we are keeping the index R implicit. Now in order to evaluate this action, we need to figure out the evolution in the radial direction. It turns out this is governed by some Hamilton-Jacobi equation. There is however an algorithmic way of evaluating this on-shell action. First of all, it will end up being local,

$$(4.52) \quad S_{\text{on-shell}} = \frac{1}{16\pi G_N} \int d^4x \sqrt{-g} \mathcal{L}$$

where we can write:

$$(4.53) \quad \mathcal{L} = \dots \frac{\mathcal{L}(4)}{\epsilon^4} + \frac{\mathcal{L}(2)}{\epsilon^2} \dots$$

note these are a bunch of divergent terms. In particular they are exactly the power-law divergences on the CFT side. So by holographic renormalization, we drop all the terms except the finite ones. So we are left with:

$$(4.54) \quad \mathcal{L} = \mathcal{L}^{(0)} + \mathcal{O}(\epsilon)$$

so this is the candidate for the renormalized Lagrangian. In even spacetime dimensions on the boundary, there is still a leftover IR divergent term. That is, there is enough room for a logarithmic divergence:

$$(4.55) \quad \mathcal{L} = \mathcal{L}^{(0)} + \mathcal{O}(\epsilon) + \log \epsilon \tilde{\mathcal{L}}^{(0)}$$

Knowing that the rescaling of the field theory corresponds to the translation in the radial direction, we can ask how the renormalized action would behave if we performed this translation in the bulk. In particular, we have a translation operator δ_D which acts as:

$$(4.56) \quad \delta_D \mathcal{L}^{(0)} \propto \tilde{\mathcal{L}}^{(0)}$$

which comes from this log divergence. So this is elementary from the bulk point of view, but from the CFT point of view, this means that the action is no longer invariant under the rescaling of the Minkowski spacetime $\vec{x} \rightarrow \lambda \vec{x}$, so there is a conformal anomaly. And if we can iteratively solve for these particular coefficients, we can find how much of a quantum anomaly is present in this classical action.

6.3. Type IIB string theory. We now present one honest list of relations for the most famous example of the AdS CFT correspondence. This example serves as a sort of sounding board for potential applications of AdS CFT correspondence. The string theory here is of type IIB on $\text{AdS}_5 \times S^5$, which will turn out to be equivalent to $\text{SU}(N)$, $\mathcal{N} = 2$ SYM in $3 + 1$ dimensions.

IIB on $\text{AdS}_5 \times S^5$,	SYM
string length: $l_s = \sqrt{\alpha'}$	N
string coupling: g_s	g_{ym}^2
$\tau_{IIB} = c_0 + i/g_s = \langle C + ie^{-\Phi} \rangle$	τ_{ym}

Note that $g_{\text{ym}}^2 = 2\pi g_s$ and

$$(4.57) \quad \tau_{\text{ym}} = \frac{\theta}{2\pi} + \frac{2\pi i}{g_{\text{ym}}^2}$$

So why is the space here $\text{AdS}_5 \times S^5$? Well IIB has C_4 , and flux $F_5 = dC_4$. So it has N units of flux through S^5 , that is

$$(4.58) \quad \int_{S^5} F_5 = N$$

Now schematically, we have the effective gravitational action:

$$(4.59) \quad S_{\text{eff}} = \frac{1}{G_N} \int \sqrt{g} R + \int F_5^2$$

Then we claim that from string theory, we get that this coefficient $1/G_N$ is the same as $1/g_s^2$, but g_s is dimensionless, whereas G_n has dimension 8 in 10 dimensions. The only parameter we have in string theory is α' , which means this is really equal to $1/(\alpha'^2 g_s^2)$. This means

$$(4.60) \quad S_{\text{eff}} \sim \frac{1}{\alpha'^2 g_s^2} L^8 + N^2$$

These two things need to balance out, which means:

$$(4.61) \quad g_s^2 \frac{L^8}{l_s^8} \sim N^2$$

So the dictionary is now very simple:

$$(4.62) \quad \boxed{\tau_{IIB} = \tau_{\text{ym}} \quad 2g_{\text{ym}}^2 N = \frac{L^4}{l_s^4}}$$

and $\lambda = g_{\text{ym}}^2 N$ is our 't Hooft coupling. So even though there is no N explicitly on the CFT time, this τ_{ym} effectively goes as N since λ is fixed. Note that on both sides, there is a $\text{SL}(2; \mathbb{Z})$ duality. This group acts on these τ s as:

$$(4.63) \quad \tau \rightarrow \frac{a\tau + b}{c\tau + d} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}(2; \mathbb{Z})$$

7. Strength of the equivalence

In general, AdS/CFT is certainly an equivalence of theories, but there is some level of uncertainty regarding the formal strength of this equivalence. In general people believe the following in a sufficiently super-symmetric setting:

CONJECTURE 1 (Strongest). *Quantum string theory, that is a string theory with:*

$$(4.64) \quad g_s \neq 0 \quad \frac{\alpha'}{L^2} \neq 0$$

is equivalent to SYM for any N and any λ .

There is the slightly more conservative version:

CONJECTURE 2 (Strong). *Classical string theory, that is a string theory with:*

$$(4.65) \quad g_s \rightarrow 0 \quad \frac{\alpha'}{L^2} \neq 0$$

is equivalent to SYM for $N \rightarrow \infty$ and λ fixed.

This weak form generally receives the most attention:

CONJECTURE 3 (Weak). *Classical (super) gravity, that is a theory of the form:*

$$(4.66) \quad g_s \rightarrow 0 \quad \frac{\alpha'}{L^2} \rightarrow 0$$

is equivalent to SYM for $N \rightarrow \infty$ and λ large.

so in this last case, we are forced into considering a strongly coupled gauge theory.

8. Anomalies

8.1. Review of anti-bracket (BRST-BV) formalism. We will mainly be using AdS/CFT later in our discussion of anomalies. Recall the BRST formalism considers some symmetry generators δ_α , fields φ^i , ghosts c^α , and antighosts $b^{4.3}$. We collectively call the fields φ^A . Let the number of fields be $N^{4.4}$. We also have the BRST charge $Q_{\text{BRST}} \equiv s$ which squares to 0. This defines a cohomology theory, which allows us to identify the Hilbert space:

$$(4.67) \quad \mathcal{H}_{\text{physical}} = H_Q(\mathcal{H}_\varphi \otimes \mathcal{H}_{b,c})$$

This formalism is however insufficient for certain theories, so we had to introduce a partner for all of these original fields. We call these anti-fields, and write them as: φ^* , c^* , b^* . Just as we did with the original fields, we call these collectively φ^{A*} where A runs from 1 to N . These anti-fields always have the opposite statistics to their original counterparts. We might now have to introduce ghosts-for-ghosts, but as long as we only introduce such fields

^{4.3} The antighosts also have an index which we omit in this brief review.

^{4.4} This is unrelated to the N from the preceding construction of large N theories.

a finite number of times, we are able to treat this sufficiently well. Now these fields φ^A and φ_{A^*} are coordinates on a $2N$ dimensional manifold. Call these coordinates $\mathcal{Z}(\varphi^A, \varphi_{A^*})$.

In the extended formalism we define this s to act as the so-called anti-bracket on any operator A :

$$(4.68) \quad sA := (A, S)$$

We should think of the anti-bracket as a sort of analogue of the Poisson bracket. Recall that:

$$(4.69) \quad s\varphi_i^* = \delta\varphi_i^* = -\frac{\delta S_0}{\delta\varphi^i}$$

We should think of S as an improved action. This starts out as a classical action, but gets additional corrections from the additional fields we introduce. We define the anti-bracket to be:

$$(4.70) \quad (A, B) = \frac{\delta^r A}{\delta\mathcal{Z}^a} \zeta^{ab} \frac{\delta^l B}{\delta\mathcal{Z}^b}$$

where ζ^{ab} is a generalized symplectic form:

$$(4.71) \quad \zeta^{ab} = \begin{pmatrix} 0 & \delta_B^A \\ -\delta_B^A & 0 \end{pmatrix}$$

Then we have

$$(4.72) \quad S = \sum_{m \geq 0} S^{(m)}$$

where m is the antighost number and we take:

$$(4.73) \quad S^{(0)} = S_{inv}(\varphi^i) \quad S^{(1)} = \varphi_i^* R_\alpha^i C^\alpha \quad \dots$$

where R_α^i is the action of the gauge transformations on the original fields. But how do we actually find S ? Essentially we just have to solve the master equation:

$$(4.74) \quad (S, S) = 0$$

This is sometimes called the Zinn-Justin equation, since they used this equation in [28] to show the renormalizability of Yang-Mills in four dimensions.^{4,5} Any solution of the master equation still has gauge invariance. In particular, this has exactly N symmetries, which we fix by choosing gauge fixing conditions $\Omega^A(\varphi, \varphi^*) = 0$. In order for these to be consistent with the structure of the master equation and $s^2 = 0$, these conditions must be “in involution.” This just means $(\Omega^A, \Omega^B) = 0$.

THEOREM 4.2. *Then under this condition $(\Omega^A, \Omega^B) = 0$ these conditions all follow from one Ψ . This is the gauge-fixing fermion.*

“PROOF”. First imagine that the $\Omega^A = 0$ conditions can be solved for φ_{A^*} . If this is not the case, we have to make a complimentary assumption. In any case, this means

$$(4.75) \quad \varphi_{A^*} - \omega_A(\varphi) = 0$$

Then:

$$(4.76) \quad (\Omega^A, \Omega^B) = \frac{\delta\omega_A}{\delta\varphi^B} - \frac{\delta\omega_B}{\delta\varphi^A} = 0 \implies \omega_A = \frac{\delta\Phi}{\delta\varphi^A}$$

^{4,5} This boils down to the fact that if we write down the most general version of Yang-Mills gauge theory in 4d, and calculate loop corrections, it turns out it is not self-contained under normalization. So to show that the theory is still renormalizable, we need anti-fields. For more details see [29].

□

So we have this solution S to the master-equation, and we can simply substitute:

$$(4.77) \quad S_{\Phi} = S \left(\varphi, \varphi^* = \frac{\delta \Phi}{\delta \varphi} \right)$$

so we have effectively eliminated the anti-fields from our formalism. Then the pragmatic BRST charge we will actual use for calculations is:

$$(4.78) \quad \bar{s}\varphi^A = (s\varphi^A) \left(\varphi, \varphi^* = \frac{\delta \Psi}{\delta \varphi} \right)$$

This means $\bar{s}S_{\Psi} = 0$, which means we can write down Feynman rules. In addition, we have that $\bar{s}^2\varphi^A$ is proportional to the field equations, which means we didn't do this all for nothing.

8.2. Anomalies. There are two types of anomalies. First, there are anomalies arising from global symmetries, which are okay, and lead to important physical effects. In this case we want to understand how significant the anomalies are, and under which symmetries they arise. The second type are anomalies arising in gauge symmetries, which invalidate quantum mechanics since we lose unitarity. This is bad, because anomalous gauge symmetric theories need to be abandoned. Mathematically, we treat them equally in the sense that we take global ones, and artificially promote these to gauge symmetries by hand. We do this since we can use the BRST formalism to treat these local symmetries. This is fine because we can always reverse this at the end of the calculation.

Consider some gauge fields A , and some effective action $\Gamma[A]$. Call the gauge transformations: $\delta_{\alpha} \equiv \mathcal{T}_{\alpha}(x)$. Under the gauge transformations,

$$(4.79) \quad \delta_{\alpha}\Gamma[A] \equiv \mathcal{T}_{\alpha} \otimes \Gamma[A] = \begin{cases} 0 & \hbar = 0 \\ G_{\alpha}[x, A] & \text{o/w} \end{cases}$$

so this is zero only in the classical limit, and otherwise we get the anomaly. Sometimes we can add local terms to Γ and remove this anomaly by hand. But what sort of forms can $G_{\alpha}[A]$ take in general? It turns out we can use the anti-bracket formalism to restrict this significantly. Recall we had the following:

$$(4.80) \quad [\delta_{\alpha}, \delta_{\beta}] = f_{\alpha\beta}{}^{\gamma} \delta_{\gamma}$$

which imply a consistency condition on G_{α} . These are known as the Wess-Zumino self consistency conditions:

$$(4.81) \quad [\delta_{\alpha}, \delta_{\beta}]\Gamma[A] = f_{\alpha\beta}{}^{\gamma} \delta_{\gamma}\Gamma[A]$$

where

$$(4.82) \quad \mathcal{T}_{\alpha}(X)G_{\beta} - \mathcal{T}_{\beta}(x)G_{\alpha} = f_{\alpha\beta}{}^{\gamma}G_{\gamma}[A]$$

Now we can simply postulate:

$$(4.83) \quad G_{\alpha}[c, A] = \int c^{\alpha}(x)G_{\alpha}[x, A]d^4x$$

So we know $s\Gamma[A] \neq 0$ and since $s^2 = 0$, we can summarize the self-consistency condition as:

$$(4.84) \quad sG[c, A] = 0$$

Therefore if we want to suggest an anomaly, then it better satisfy this condition.

So now it is easy to classify the anomalies. We can only have G s such that:

- (1) $\text{gh}\#(G) = 1$
- (2) $sG = 0$
- (3) $G + s\Gamma_1 \cong G$

where Γ_1 is local.

8.3. Schwinger terms. Consider equal time commutation relations of conserved currents.

$$(4.85) \quad [J_\alpha^0(\vec{x}, t), J_\beta^0(\vec{y}, t)] = f_{\alpha\beta}{}^\gamma J_\gamma^0(\vec{x}, t) \delta^{d-1}(\vec{x} - \vec{y}) + S_{\alpha\beta}(\vec{x}, \vec{y}, t)$$

Now in analogy with the treatment of anomalies, we can define

$$(4.86) \quad S = \int d^{d-1}\vec{x} d^{d-1}\vec{y} c^\alpha(x) c^\beta(y) S_{\alpha\beta}(\vec{x}, \vec{y}, t)$$

This object has ghost number 2. Then the claim is, that this object is the second cohomology with respect to the BRST charge, which classifies the possible Schwinger terms:

- (1) $sS = 0$ by virtue of the Jacobi identity.
- (2) If $S = s\mathcal{J}$, then \mathcal{J} has to be of the form $\int c^\alpha \mathcal{J}_\alpha$ for some \mathcal{J}_α , which can be used to define shifted currents:

$$(4.87) \quad \tilde{J}_\alpha^0 = J_\alpha^0 \pm \mathcal{J}_\alpha$$

which satisfy the expected commutation relation without the Schwinger anomaly.

8.4. Examples. We first consider Weyl anomalies in relativistic systems. These are associated with transformations of the form:

$$(4.88) \quad g_{\mu\nu}(x) \rightarrow \Omega^2(x) g_{\mu\nu}$$

Note that if $g_{\mu\nu} = \eta_{\mu\nu}$, we get Ω^2 is just some rescaling λ , so these global conformal transformations are related to RG flows. That is, this is some sort of local version of conformal transformations. So if we have some action $\Gamma[g_{\mu\nu}]$, we might have an anomaly under these so-called Weyl transformations.

EXAMPLE 4.2. In $1+1$ dimensions, the anomaly will look like:

$$(4.89) \quad \mathcal{A} \propto G_\alpha[g_{\mu\nu}] c \int d^{1+1}x \sqrt{-g} R \quad G = \frac{c}{\#} \int \mathbf{c} \sqrt{-g} R$$

where \mathbf{c}^α is the ghost, c is the central charge, and R is the Ricci scalar. Recall $\sqrt{-g}R$ is a topological invariant.

In general, any $d-1, 1$ theory has a conformal $\text{SO}(d, 2)$ symmetry. So for $d=2$ we get

$$(4.90) \quad \text{SO}(2, 2) \cong \text{SO}(2, 1) \oplus \text{SO}(2, 1)$$

In addition to this, CFTs in two dimensions satisfy the so-called Virasoro algebra:

$$(4.91) \quad [L_m, L_n] = (m-n)L_{m+n} + \frac{c}{12}m(m^2-1)\delta_{m+n,0}$$

In a unitary CFT, one can show that $c \geq 0$, and $c=0$ only for the trivial theory, which has a Hilbert space with only the vacuum state. This c is useful because it is a quantitative measure of the number of degrees of freedom. Note that it is normalized such that $c=1$ for

$$(4.92) \quad S = \frac{1}{2} \int (\partial\varphi)^2$$

THEOREM 4.3 (Zamolodchikov). *In 1 + 1 dimension, there exists a function $C(g_i, \mu)$, where μ is the RG scale, such that C decreases monotonically along the RG flow (from the UV to the IR), and at any RG fixed point g_* we have $C(g_*) = c$.*

We now redo the previous example in $2d + 1$ dimensions, and we find out $H_s^1 = 0$ so there are no Weyl anomalies in any odd spacetime dimension.

EXAMPLE 4.3. In $3 + 1$ dimensions, we have

$$(4.93) \quad \mathcal{A} \sim a \int E_4 + cI$$

where E_4 is the Euler number density:

$$(4.94) \quad E_4 \sim \sqrt{-g} (R_{\mu\nu\rho\sigma}^2 - 4R_{\mu\nu}^2 + R^2)$$

So we have two candidate central charges a and c . This a is a closer analogue to the central charge in $1 + 1$ dimensions. In $3 + 1$ dimensions we can define the Weyl-tensor, which we were not able to do in $1 + 1$ dimensions:

$$(4.95) \quad W_{\nu\sigma\rho}^\mu = R_{\nu\sigma\rho}^\mu - \text{traces}$$

Then we can write this I factor as:

$$(4.96) \quad I \propto \int \sqrt{-g} W$$

Then we have the following theorem:

THEOREM 4.4 (*a*-theorem). *This a is the measure of numbers of degrees of freedom in the field theory.*

This was proposed by J. Cardy in 1988, but was “proven” by Komargodski and Schwimmer in 2011. They believe to have a nonperturbative proof [15].

REMARK 4.3. For CFTs in $3 + 1$ dimensions which have purely gravitational dual, it has been proven that automatically $a = c$.

EXAMPLE 4.4. In $5 + 1$ dimensions, we have

$$(4.97) \quad \mathcal{A} \sim a \int E_6 + \sum_{i=1}^3 z_i I_i(W)$$

where a, c_i are the 4 “central charges.”^{4.6}

Note that we can repeat this analysis in the context of nonrelativistic QFTs. The analogue of the Weyl translation then becomes:

$$(4.98) \quad t \rightarrow \lambda^z t \qquad x^i \qquad \rightarrow \lambda x^i$$

where $z > 0$ is the “dynamical critical exponent” which is an important observable quantity associated to an RG fixed point. If $z = 1$, we are back to the relativistic case, and $z = 2$ gives us Galilean invariant theories. So this is just a generalization of the Weyl translation we saw before. Some of the statements which we are used to in relativistic field theories will now not be true here.

^{4.6} We put this in quotes, because there is no sense in which any of these are the analogue of the actual central charge in the sense of the Virasoro algebra.

EXAMPLE 4.5. Consider a theory in $2 + 1$ dimensions for $z = 2$. Define Weyl transformations by:

$$(4.99) \quad g_{\mu\nu} = \begin{pmatrix} -N^2 + g^{ij}N_iN_j & N_i \\ N_i & g_{ij} \end{pmatrix}$$

this form of the metric is called the ADM decomposition. So g_{ij} is like the spatial portion of the metric, N_i is like the shift vector, and N is the lapse function.

Then in this language, the anisotropic Weyl transformations are given by:

$$(4.100) \quad g_{ij} \rightarrow \Omega^2(t, \vec{x}) g_{ij} \quad N_i \rightarrow \Omega^2 N_i \quad N \rightarrow \Omega^z N$$

Now we can write:

$$\begin{aligned} \mathcal{A} = c_K \int d^2x dt \sqrt{g} N \left[K_{ij} K^{ij} - \frac{1}{2} K^2 \right] \\ + c_V \int d^3x dt \sqrt{g} N \left[R - \left(\frac{\nabla_i N}{N} \right)^2 + \frac{\nabla N}{N} \right]^2 \end{aligned}$$

So we have two ‘‘central charges,’’ a kinetic one c_K , and a potential one c_V . Note K_{ij} is the extrinsic curvature:

$$(4.101) \quad K_{ij} = \frac{1}{2N} (\dot{g}_{ij} - \nabla_i N_j - \nabla_j N_i)$$

What about isometries in one higher dimension? We will get something similar to our treatment of AdS. We get something like:

$$(4.102) \quad ds^2 = -\frac{dt^2}{u^{2z}} + \frac{d\vec{x}^2}{u^2} + \frac{du^2}{u^2}$$

Note that for $z = 1$, this is not an AdS. This is often times referred to as Lifshitz spacetime.

So we can start with GR (i.e. $\Lambda < 0$) and add whatever $I_{\mu\nu}$. The problem is, very often these tensors are not self-consistent. So we have that for nonrelativistic CFT, having a relativistic dual, means $c_K \neq 0$ and $c_V = 0$. Alternatively, we can start with nonrelativistic gravity with no matter. In this case with nonrelativistic gravity duals, we have $c_K \neq 0$ and $c_V \neq 0$.

8.5. The anti-bracket in practice. The following claim is often made about the SM or even generic YM:

CLAIM 4.1. Chiral anomalies are one-loop exact.

To show this, one must invoke the full structure of the Zinn-Justin equation. That is, write

$$(4.103) \quad \Gamma = S + \Gamma_1 + \Gamma_2 + \dots$$

then if the theory is free of anomalies, it must satisfy the master equation $(\Gamma, \Gamma) = 0$. If there are only one-loop anomalies, we have $(S, \Gamma_1) = G_1 \neq 0$. But then the anti-bracket has trivial BRST cohomology for this particular anomaly at all orders $n \geq 2$.

So consider some arbitrary non-abelian gauge theory. We will now concern ourselves over the gauge anomalies here. We treat this using the so-called descent equations (or Stora-Zumino descent equations). The idea is to move from our generic number of dimensions $d = 2n$ to some higher dimensional scenario $2n + 2$:

$$(4.104) \quad \Omega_{2n+2} = \text{Tr} \underbrace{(F \wedge \dots \wedge F)}_{n+1}$$

In a formal sense, we clearly have $s\Omega_{2n+2}^{(0)} = 0$. We know that since this is BRST closed, locally it will be BRST exact. Then by definition, it also satisfies the local condition that it is the total derivative of some $2n + 1$ differential form:

$$(4.105) \quad \Omega_{2n+2}^{(0)} = d\Omega_{2n+1}^{(0)}$$

We also know that $sd + ds = 0$, so now we can write:

$$(4.106) \quad d(s\Omega_{2n+1}) = -sd(\Omega_{2n+1}) = -s^2(\dots) = 0$$

This means that $s\Omega_{2n+1}$ is the derivative of some object of ghost number 1:

$$(4.107) \quad s\Omega_{2n+1} = d\Omega_{2n}^{(1)} \quad G = \int \Omega_{2n}^{(1)}$$

We could push this even further to get that

$$(4.108) \quad d(s\Omega_{2n}^{(1)}) = 0$$

which implies that

$$(4.109) \quad s\Omega_{2n}^{(1)} = d\Omega_{2n-1}^{(2)}$$

so these $\Omega^{(n)}$ objects we are getting are the analogues of the $S_{\text{Schwinger}}$ objects from before.

Effective field theory of gravity

1. Preliminaries

We know that gravity is well described by the Einstein-Hilbert action:

$$(5.1) \quad S = \frac{1}{16\pi G_N} \int d^D x \sqrt{-g} (R) \quad G_N \sim \frac{1}{M_{\text{pl}}^{D-2}}$$

However, we also know that this theory is non-renormalizable, since $[G_n] < 0$ for $D > 2$. UV completion suggests a breakdown at $E \sim M_{\text{pl}}$. But there is also an IR breakdown so we can also add a nontrivial constant term:

$$(5.2) \quad S = \frac{1}{16\pi G_N} \int d^D x \sqrt{-g} (R - 2\Lambda)$$

Note that $[\Lambda] = 2$, since it sits on the same footing as R , which has dimension 2. If we take the cosmological constant problem seriously, we see that there should be a breakdown at ~ 1 eV. So if we want to embed this into a quantum field theory, we are facing multiple problems.

2. UV-IR connection?

Phenomenologically, we see ourselves living in a universe where Λ is much much smaller than M_{pl} . So maybe we should just interpret this as indicating a gravitational connection. More pragmatically, we can consider the existence of black holes, which implies the holographic principle. The scaling of S with the area \mathcal{A} suggests the scaling of the number of degrees of freedom. Cosmology also suggests this connection. This is driven by the dramatic time dependence in the evolution of the universe.

Note that we can have gravity with $\Lambda < 0$, $\Lambda > 0$, and flat spacetime $\Lambda = 0$. It is known that holography displays AdS/CFT behaviour for $\Lambda < 0$. Here we have consistency with SUSY, and string theory, so there is no cosmological constant problem in this highly SUSY case. There is a UV-IR connection in this case, but this is still an open problem

For $\Lambda > 0$, this is our real universe, we have no AdS/CFT, but we might have some sort of de Sitter space CFT duality. There is also no SUSY consistency here. String theory also does a bad job of describing this, for reasons that are still somewhat mysterious. Therefore we have to resort to EFT techniques. In this universe we have many interesting results such as the horizon problem, and the flatness problem. A priori, we don't know what solutions these might have, but inflation might provide an explanation for this.

3. Review of EFT

The generic idea of EFT is that we don't need to understand everything up to arbitrarily high energy to form an understanding of low-energy physics. In addition, the physics of the low energy should be independent of the high energy physics. That is, we are essentially

But what about corrections at tree level? For gravity, we don't a priori know that this makes sense for this non-renormalizable theory, but it turns out there are meaningful loop corrections. In order to determine which questions are meaningful and which ones aren't we first rename the fields a bit:

$$(5.10) \quad \sigma + i\vec{\tau} \cdot \vec{\pi} (v + \sigma') e^{i\vec{\tau} \cdot \vec{\pi}'/v} = v + \sigma' + i\vec{\tau} \cdot \vec{\pi}' + \dots = v + \sigma + i\vec{\tau} \cdot \vec{\pi} + \dots$$

which should not change the physics. This gives us:

$$(5.11) \quad S = \frac{1}{4} \int d^4x (v + \sigma)^2 \text{Tr} (\partial_\mu U \partial^\mu U^\dagger) + S[\sigma]$$

Now we want to integrate out the heavy mode σ . At leading order, ignoring σ give us a good approximation, which gives us the EFT from before: $S_{\text{eff}} \equiv S_{\text{eff}}^{(0)}$. To consider the next order, we can no longer ignore σ . This is just a simple Gaussian integral, which yields:

$$(5.12) \quad S_{\text{eff}}^{(0)} + \frac{v^2}{8m_\sigma^2} \int \text{Tr} (\partial_\mu U \partial^\mu U^\dagger)$$

Now how do we treat loops and renormalization? More generally, the local Lagrangian for S_{eff} , with unknown coefficients, in energy expansion yields q^2/v^2 . Well we know that the scale we expect things to break down is given by:

$$(5.13) \quad S_{\text{eff}} = \frac{M_{\text{pl}}^2}{4} \text{Tr} (\partial_\mu U \partial^\mu U^\dagger) + l_1 (\text{Tr} (\partial_\mu U \partial^\mu U^\dagger))^2 + l_2 \text{Tr} (\partial_\mu U \partial^\mu U) \text{Tr} (\partial^\mu \partial_\nu U) + \mathcal{O}(\partial^6)$$

Now the matching conditions predict:

$$(5.14) \quad l_1 = \frac{v^2}{8m_\sigma^2} \quad l_2 = 0$$

but the renormalized one-loop corrected values will be:

$$(5.15) \quad l_1^R = \frac{v^2}{8m_\sigma^2} + \frac{1}{384\pi^2} \left[\log \frac{m_\sigma^2}{\mu^2} - \frac{35}{6} \right] \quad l_2^R = \frac{1}{192\pi^2} \left[\log \frac{m_\sigma^2}{\mu^2} - \frac{11}{6} \right]$$

4.2. Non-linear σ -model. Now we can apply this to the non-linear σ model, and we won't be able to do matching, but we can follow roughly the same steps. The action here will be:

$$(5.16) \quad S_{gr} = \int d^4x \sqrt{-g} \left\{ \frac{1}{16\pi G_n} (R - 2\Lambda) + c_1 R^2 + c_2 R_{\mu\nu} R^{\mu\nu} + c_3 R_{\mu\nu\sigma\rho} R^{\mu\nu\sigma\rho} + c_1^{(3)} R^3 + c_1^{(4)} R^4 + \dots \right\}$$

In the case of the NL σ M, we have the following:

THEOREM 5.1 (Weinberg). *This systematic expansion in the powers of momentum makes sense.*

EXAMPLE 5.1. We can calculate graviton-graviton scattering with this approach.

$$(5.17) \quad \mathcal{A}(++,++) = \frac{i k^2 s^3}{4 tu} \left\{ 1 + \frac{k^2 stu}{\#} \left[\frac{\log -u}{st} + \dots \right] + \dots \right\}$$

We also get quantum corrections to the potential:

$$(5.18) \quad V(r) = -\frac{GMm}{r} \left[1 + \frac{G(M+m)}{rc^2} + \frac{41G\hbar}{10\pi r^2 c^3} + \dots \right]$$

The interesting features here are more visible in momentum space. We get $1/r \sim 1/q^2$, and then the quantum corrections go like:

$$(5.19) \quad \frac{1}{r^2} \sim \frac{1}{q^2} \sqrt{q^2} \qquad \frac{1}{r^3} \sim \frac{1}{q^2} q^2 \log q$$

so we have the expected propagator, multiplied by something non-analytic in q . This is a sign that this came from the 1-loop correction. On the other hand, if we had a local term such as $\delta^3(\vec{x})$ we get:

$$(5.20) \quad \delta^3(\vec{x}) \sim \frac{1}{q^2} q^2$$

so terms like this are outside of the scope of this technique.

For more information on this approach, see [8, 9].

4.3. Technical naturalness. This isn't really a great EFT (yet). This is a result of technical naturalness. This is an idea which first came from 't Hooft. The initial concept was that microscopic physics implies macroscopic physics. That is, the theory flows from the UV to the IR. The mathematical statement here is that the parameters α in S can be small, as long as $\alpha = 0$ implies enhanced symmetry. This means, that assuming technical naturalness, Λ_c cannot be "small" since setting $\Lambda_c = 0$ doesn't give us any enhanced symmetry. Recall that for $\Lambda_c < 0$, we have the AdS/CFT correspondence, and an $SO(3, 2)$. For $\Lambda_c = 0$, we have the inhomogeneous $SO(3, 1)$. For $\Lambda_c > 0$, we have $SO(4, 1)$ in de Sitter space. In all of these cases, $\dim G = 10$. This means $\Lambda_c = 0$ doesn't enhance any symmetries, so not only can it not be small, but in fact $\Lambda_c \sim \Lambda^2$. So this theory doesn't break down at energies smaller or equal to the order of M_{pl} , but rather $E < \sqrt{\Lambda_c}$.

EXAMPLE 5.2. Consider the following scalar field theory:

$$(5.21) \quad S = \frac{1}{2} \int d^4x \left((\partial\varphi)^2 - m^2\varphi^2 - \frac{\lambda}{12}\varphi^2 \right)$$

This is already self interaction without adding anything else. But what kinds of symmetries do we have here? Well we have the trivial shift symmetry $\varphi \mapsto \varphi + c$. Is this theory invariant? Well we see an issue with the second two terms, but when these are 0, this means we have an enhanced symmetry. Meaning we can self consistently make $m^2 \sim \epsilon M_{pl}^2$ and $\lambda \sim \epsilon$ for $\epsilon \ll 1$ by the postulate of technical naturalness.

For more on EFT of gravity see [4, 5].

5. Effective field theory of inflation

5.1. Cosmological setup. We know that we live in a cosmological spacetime, and by the Copernican principal that we aren't in any sort of special place, we will assume that we have a so-called FLRW^{5.1} metric. These metrics are of the form:

$$(5.22) \quad ds^2 = -dt^2 + a^2(t) \left[\frac{dr^2}{1 - kr^2} + r^2 d\Omega^2 \right]$$

where

$$(5.23) \quad k = \begin{cases} 0 & \mathbb{R}^3 \\ 1 & S^3 \\ -1 & H^3 \end{cases}$$

^{5.1} This stands for Friedmann - Lemaître - Robertson - Walker. This is sometimes shortened to FRW.

Now we can define a so-called Hubble parameter:

$$(5.24) \quad H(t) = \partial_t \log a(t) = \frac{\dot{a}(t)}{a(t)}$$

As for the dynamics, we can either take our favorite EFT of gravity, or the Einstein equations. We will take the latter. In this context, this gives us the Friedmann equations:

$$(5.25) \quad H^2 = \frac{1}{3M_{pl}^2} \rho \quad \dot{H} + H^2 = -\frac{1}{6M_{pl}^2} (\rho + 3p)$$

It is often useful to introduce conformal time:

$$(5.26) \quad d\tau = \frac{dt}{a(t)}$$

In this language the metric becomes:

$$(5.27) \quad ds^2 = a^2(\tau) [-d\tau^2 + dr^2 + r^2 d\Omega^2]$$

Now in order to talk about inflation, we want to consider the causal structure on these cosmologies. In conformal time, geodesics are straight lines, and at angles $\pi/4$ in the (τ, r) plane. We want to calculate the maximum distance a photon can travel in a given time-span:

$$(5.28) \quad \Delta r = \Delta\tau = \int_{t_0}^{t_f} \frac{dt'}{a(t')}$$

Then if we take t_i to be the origin of the universe, we have $\Delta r_{\max}(t)$ is the *comoving particle horizon*. We can rewrite this as:

$$(5.29) \quad \tau = \int \frac{dt}{a(t)} = \int \underbrace{\frac{1}{aH}}_{\text{comoving Hubble radius}} d \log a$$

EXAMPLE 5.3. Consider a system with matter, and $w = p/\rho$. We can then solve the Friedmann equations to get:

$$(5.30) \quad \frac{1}{aH} \propto a^{(1+3w)/2}$$

In the GR literature, the case that $1 + 3w > 0$ is known as the strong energy condition (SEC).

5.2. Inflation. Most people who are advocates of inflation, cite that this “solves” the “horizon” and “flatness” problem. The horizon problem is, roughly, that if we have the SEC, this disagrees with the observed homogeneity of the cosmological microwave background. That is, if we have some finite initial conformal time t_i , when we measure background at some early conformal time τ_{rec} , there isn’t enough time for photons to move in such a way that allows for the homogeneity that we observe. See fig. 1 for a diagram of this. The solution to this is effectively letting the universe infinitely long before this recorded time, so this shaded region covers the full region. This means:

$$(5.31) \quad \frac{d}{dt} \left(\frac{1}{aH} \right) < 0$$

We can turn this apparent contraction around, and end up with a different diagram. So the period where this is satisfied can be called the *inflationary time period*. Then if we have $w < -1/3$, we might be overturning the consequence of the SEC, and we might end up solving this inequality. For example, we might have a system with a matter source, and Λ_c with $w = -1$.

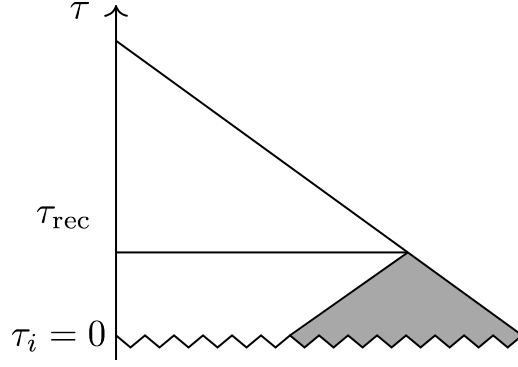


FIGURE 1. Conformal diagram showing the horizon problem.

The flatness problem is, briefly, that if we start with a somewhat curved spacetime, if we just apply FRW without inflation, spacetime becomes very curved. With inflation, we get a spacetime which is roughly as flat as we seem to observe it.

In order for this to be consistent, we need

$$(5.32) \quad \epsilon := -\frac{\dot{H}}{H^2} = \frac{3}{2} \left(1 + \frac{p}{\rho}\right) < 1 \quad |\eta| := \frac{|\dot{\epsilon}|}{H\epsilon} \ll 1$$

the second condition assures that inflation lasts long enough.

We now consider an example called slow roll inflation. Consider some theory:

$$(5.33) \quad S = \int d^4x \sqrt{-g} \left(\frac{M_{pl}}{2} R - \frac{1}{2} g^{\mu\nu} \partial_\mu \varphi \partial_\nu \varphi - V(\varphi) \right)$$

where the field φ is called an inflaton. So what conditions do we have on $V(\varphi)$? Well we saw the two conditions on inflation already, so just looking at the EOM for this system we can just calculate:

$$(5.34) \quad \epsilon = \frac{\dot{\varphi}^2}{2M_{pl}^2 H^2} \quad \eta := 2(\epsilon - \delta) \quad \delta = -\frac{\ddot{\varphi}}{H\dot{\varphi}}$$

This leads to the so-called eta problem which has several different incarnations. Note that we are already fine tuning so the constant piece is absent, which is already somewhat unsatisfying. In addition, the potential $V(\varphi)$, starts as $m^2\varphi/2$, and then it turns out that the quantum correction of η is given by:

$$(5.35) \quad \Delta\eta \approx M_{pl}^2 \frac{\Delta V''}{V} \sim \frac{\Lambda}{H^2}$$

so the punchline is, that if we have a UV cut-off in an EFT of inflation, then we need to choose this cut-off, and if this cut-off is the plank scale, this is a huge contribution, much larger than 1. We might then think the cut-off should only be H . But we can't solve the CC problem, so maybe we shouldn't even be allowed to do this. In any case, this still won't really help us, because this still just gives a correction to η of order 1. There is no symmetry gained by making $\eta = 0$, so in the least exciting possible scenario, this correction is still of order 1.

There is also a so-called *strong eta problem*, which considers an EFT with an action of the form:

$$(5.36) \quad S = \dots \int \frac{c_1}{\Lambda^8} \varphi^2 (\partial\varphi)^2 R^2 + \dots$$

then the only way out is to postulate the shift symmetry, but then coupling to gravity breaks this symmetry. So this strong problem says the inflationary physics is sensitive to an infinite collection of higher dimensional operators c_i . So an infinite sequence of these will substantially influence inflation, and we therefore have to complete an infinite sequence of fine-tuning.

For more information on EFT of inflation, see **[2, 3, 6]**.

Applications of QFT to finance

Much of the content and figures from this section are from [14].

1. Fluctuations of financial assets

Denote the price of a stock, or some other asset, by $S(t)$. This is a complex thing to follow, but over a long period of time, we start to see some trends. In particular, this is relatively well approximated by exponentials as plotted in fig. 1. The fluctuations of this $S(t)$ have an average width, which we will call the *volatility of the market*, written as σ . Over long-times, that is, over a day or month, the volatility σ behaves stochastically as illustrated in fig. 2. The volatilities approximately follow a Gamma distribution as illustrated in fig. 3. The normalized log-normal distribution has the form:

$$(6.1) \quad D^{\text{log-norm}}(z) = (2\pi\sigma^2 z^2)^{-1/2} \exp\left(-(\log z - \mu)^2 / 2\sigma^2\right) .$$

1.1. Harmonic approximations. The fluctuations of S are roughly given by the following expression:

$$(6.2) \quad \frac{\dot{S}(t)}{S(t)} = r_S + \eta(t)$$

where r_S is the growth rate, and $\eta(t)$ is a white noise function given by the following:

$$(6.3) \quad \langle \eta(t) \rangle = 0 \quad \langle \eta(t) \eta(t') \rangle = \sigma^2 \delta(t - t')$$

Note that we will also call this $v := \sigma^2$, the variance.

We call $dS(t) / S(t)$ to be the *return* of the asset. Financial data is given with discrete time intervals δt , so often times this is used instead of the infinitesimal dt . Define the

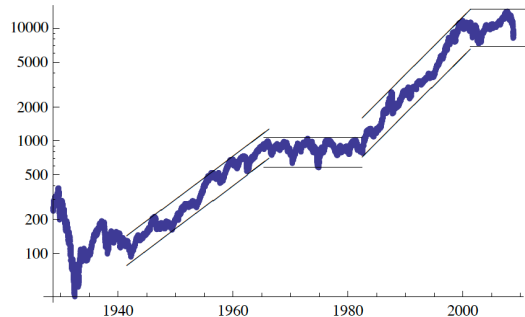


FIGURE 1. Logarithm of the Dow Jones industrial index over a range of 80 years.

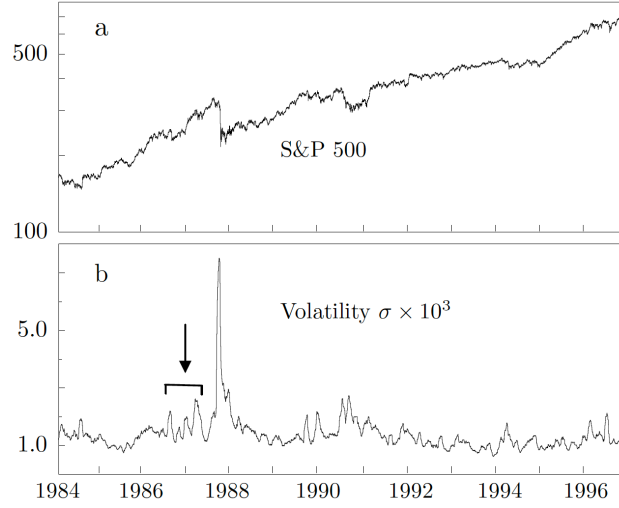


FIGURE 2. (a) Index S&P 500 for over 13 years, recorded every minute. (b) Volatility every 30 minutes.

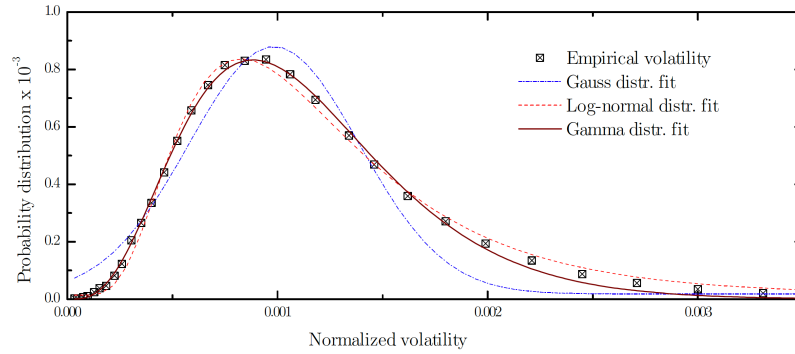


FIGURE 3. Comparison of best Gaussian, log-normal, and Gamma distribution to volatilities over 300 minutes. The normalized log-normal distribution has the form in (6.1). We will revisit the Gamma distribution.

following:

$$(6.4) \quad x(t) := \log S(t)$$

This implies a stochastic differential equation for linear growth:

$$(6.5) \quad \dot{x}(t) = \frac{\dot{S}}{S} - \frac{1}{2}\sigma^2 + r_x + \eta(t)$$

where we have taken

$$(6.6) \quad r_x := r_s - \frac{1}{2}\sigma^2$$

to be the *drift* of the process. But where does this $\sigma^2/2$ come from?

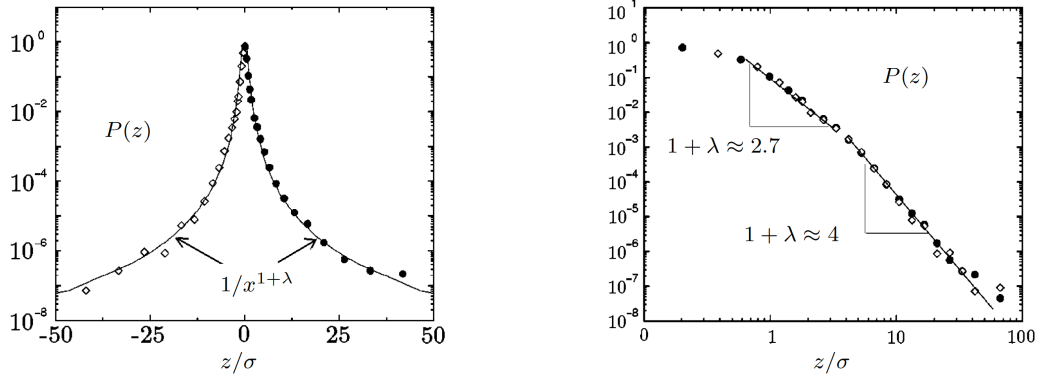


FIGURE 4. Left: Lévy tails of the S&P 500 index (1 minute log-returns) plotted against z/δ . Right: Double-logarithmic plot exhibiting power-like tail regions of the S&P 500 index (1 minute log-returns).

For functions of a stochastic variable $x(t)$ we can take the following formal expansion in dt :

$$(6.7) \quad dx(t) = \frac{dx}{dS} dS(t) + \frac{1}{2} \frac{d^2x}{dS^2} dS^2(t) + \dots$$

$$(6.8) \quad = \frac{\dot{S}(t)}{S(t)} dt - \frac{1}{2} \left[\frac{\dot{S}(t)}{S(t)} \right]^2 dt^2 + \dots$$

now we can take $\dot{x}^2 dt \mapsto \langle \dot{x}^2 \rangle dt = \sigma^2$ so we have

$$(6.9) \quad \left[\frac{\dot{S}(t)}{S(t)} \right]^2 dt \mapsto \dot{x}^2(t) dt = \sigma^2$$

1.2. Approximations by distributions. This description of the logarithms of the stock prices by Gaussian fluctuations around a linear trend is clearly only a rough approximation. At small time intervals, such as every minute, the logarithms of the stock prices are well approximated by Gaussians. However, for rare events, there is a much higher probability than in Gaussians. Such distributions are observed to have *heavy tails* in comparison with the light tails of Gaussian distributions.

EXAMPLE 6.1. In an attempt to fix this heavy-tail issue, we might consider the Fourier transform:

$$(6.10) \quad \tilde{L}_{\sigma^2}^\lambda(z) := \int_{-\infty}^{\infty} \frac{dp}{2\pi} e^{ipz} L_{\sigma^2}^\lambda(p)$$

where we take:

$$(6.11) \quad L_{\sigma^2}^\lambda(p) := \exp \left[-(\sigma^2 p^2)^{\lambda/2} / 2 \right]$$

Phenomenologically, this λ is taken between 1.2 – 1.5. One might consider whether this λ could be a critical exponent of some RG fixed point. Experimentally, these are referred to as Lévy tails. See fig. 4

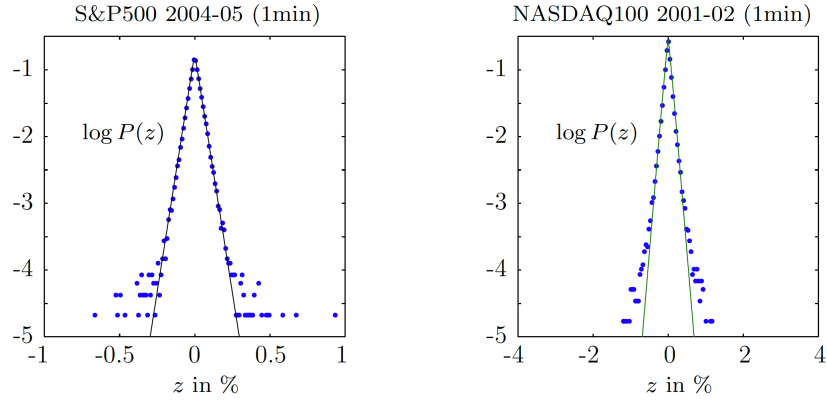


FIGURE 5. Boltzmann distribution of S&P 500 and NASDAQ 100 high-frequency log-returns recorded by the minute.

Before listing additional examples of distributions, note that for any distribution $\tilde{D}(z)$, we can define some useful objects. Take the Fourier decomposition:

$$(6.12) \quad \tilde{D}(z) \int \frac{dp}{2\pi} e^{ipz} D(p)$$

and we can define the Hamiltonian to be $H(p)$ such that:

$$(6.13) \quad D(p) \equiv e^{-H(p)}$$

Equivalently, we might take the following as the definition:

$$(6.14) \quad e^{-H(p)} \equiv \langle e^{-ipz} \rangle$$

Note that now we can identify the Lévy distributions as having a Hamiltonian given by

$$(6.15) \quad H(p) = \frac{1}{2} (\sigma^2 p^2)^{\lambda/2}$$

EXAMPLE 6.2. We now see the Gamma distribution mentioned earlier:

$$(6.16) \quad \tilde{D}_{\mu,v}^{\text{Gamma}}(z) = \frac{1}{\Gamma(v)} \mu^v z^{v-1} e^{-\mu z}$$

This has the following Hamiltonian:

$$(6.17) \quad H(p) = v \log \left(1 - \frac{ip}{\mu} \right)$$

1.3. Boltzmann distribution. The first question a physicist might ask about all this, is if we can get away with modelling this system with a Boltzmann distribution. That is, does the system have a characterising temperature? As it turns out, this point of view leads to a high-performing model. The data for the S&P 500 and NASDAQ 100 are displayed in fig. 5.

The data are fitted by the distribution:

$$(6.18) \quad \tilde{B}(z) = \frac{1}{2T} e^{-|z|/T}$$

It is evident that only a small portion of events (for large $|z|$) do we get a poor fit. These events display this heavy-tail behaviour. The temperature associated with the stock-market

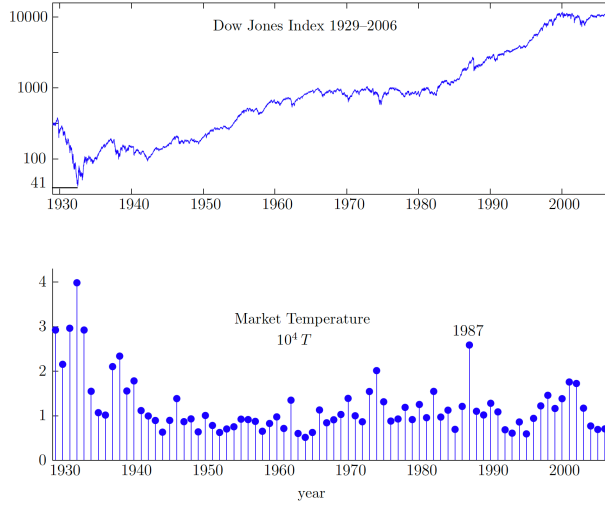


FIGURE 6. Dow Jones index over 78 years (1929 – 2006) and the annual market temperature, which is remarkably uniform, except in the 1930's, in the beginning of the great depression. Another high extreme temperature occurred in the crash year 1987.

changes relatively slowly with respect to the socio-political climate, and is high in times of crisis. We can see in fig. 6 that crashes give us high temperatures.

There are many other distributions which can do a relatively good job of fitting the available economic data, but we omit any discussion of these. Many of these distributions seem ad hoc. One might wonder if a physicist can use the same intuition that drives the consideration of new physical theories to consider intuitive models rather than models with arbitrary convenient ansatz. We now develop the path integral formalism in this context.

2. Path integrals

Consider the following stochastic differential equation for the logarithms of S :

$$(6.19) \quad \dot{x}(t) = r_x + \eta'(t) \quad \langle \eta'(t) \rangle$$

This η is the noise variable, and is arbitrarily distributed. The constant drift r_x is only uniquely defined when the average $\langle \eta \rangle = 0$. This isn't generally true, since plenty of the distributions used in practice have nonzero expectation. In any case, we redefine the constant in the expansion of the Hamiltonian so we can just write $\dot{x}(t) = \eta(t)$. Now the probability distribution of the endpoints $x_b = x(t_b)$ for the paths starting at a certain initial point $x_a = x(t_a)$ can be written as a path integral:

$$(6.20) \quad P(x_b t_b | x_a t_a) = \int \mathcal{D}\eta(t) \int_{x_a}^{x_b} \mathcal{D}x(t) \exp\left(-\int_{t_a}^{t_b} dt \tilde{H}(\eta(t))\right) \delta[\dot{x} - \eta]$$

This $\tilde{H}(\eta)$ is defined to be:

$$(6.21) \quad \tilde{H}(\eta) = -\log \tilde{D}(\eta)$$

In the mathematical literature, the measure in the path integral over the noise

$$(6.22) \quad \mathcal{D}\mu \equiv \mathcal{D}\eta P[\eta] = \mathcal{D}\eta \exp\left(-\int_{t_a}^{t_b} dt \tilde{H}(\eta(t))\right)$$

of the probability distributions is called the *measure of the process* $\dot{x}(t)$. The path integral:

$$(6.23) \quad \int_{x_a}^{x_b} \mathcal{D}x \delta[x - \eta]$$

is called the *filter* which determines the distribution of x_b at time t_b for all paths $x(t)$ starting at $x_a = x(t_a)$.

The correlation functions of the noise variable $\eta(T)$ in eq. (6.20) are given by a straightforward functional generalization of what we have already seen. In particular, we express the noise distribution $P[\eta]$ from eq. (6.20) as a Fourier path integral

$$(6.24) \quad P[\eta] = \exp\left(-\int dt \tilde{H}\right) = \int \frac{\mathcal{D}p(t)}{2\pi} \exp\left(\int dt [ip(t)\eta(t) - H(p(t))]\right)$$

2.1. Time evolution of distribution. The δ function in eq. (6.20) may be represented by a Fourier integral leading to the path integral

$$(6.25) \quad P(b|a) = \int \mathcal{D}\eta \int \mathcal{D}x \int \mathcal{D}p \exp\left(\int i\dot{x} - ip\eta - \tilde{H}(\eta(t))\right)$$

$$(6.26) \quad = \int \mathcal{D}x \int \frac{\mathcal{D}p}{2\pi} \exp\left(\int dt (ip\dot{x} - H(p(t)))\right)$$

$$(6.27) \quad = \int_{-\infty}^{\infty} \frac{dp}{(2\pi)} \exp(ip(x_b - x_a) - (t_b - t_a)H(p))$$

where we have integrated out the noise variable, and integrated over all $x(t)$ with fixed endpoints. We simply write this as:

$$(6.28) \quad P(x, t) = \int_{-\infty}^{\infty} \frac{dp}{2\pi} \exp(ipx - tH(p))$$

Note that as expected, this satisfies the Markovian property:

$$(6.29) \quad P(b|a) = \int dx_c P(b|c) P(c|a)$$

as well as a Fokker-Planck-type equation

$$(6.30) \quad \partial_t P(b|a) = -H(-i\partial_x) P(b|a)$$

For more on this subject see [14, 25].

Bibliography

- [1] L. Baulieu and I. M. Singer, *The topological sigma model*, Comm. Math. Phys. **125** (1989), no. 2, 227–237.
- [2] D. Baumann, *TASI Lectures on Inflation*, ArXiv e-prints (July 2009), available at [arXiv:0907.5424](https://arxiv.org/abs/0907.5424).
- [3] D. Baumann and L. McAllister, *Inflation and String Theory*, 2015.
- [4] C. P. Burgess, *Quantum Gravity in Everyday Life: General Relativity as an Effective Field Theory*, Living Reviews in Relativity **7** (April 2004), 5, available at [arXiv:gr-qc/0311082](https://arxiv.org/abs/gr-qc/0311082).
- [5] ———, *Effective Theories and Modifications of Gravity*, ArXiv e-prints (December 2009), available at [arXiv:0912.4295](https://arxiv.org/abs/0912.4295).
- [6] ———, *Intro to Effective Field Theories and Inflation*, ArXiv e-prints (November 2017), available at [arXiv:1711.10592](https://arxiv.org/abs/1711.10592).
- [7] S. Coleman, *Aspects of symmetry: Selected erice lectures*, Cambridge University Press, 1988.
- [8] J. F. Donoghue, *The effective field theory treatment of quantum gravity*, American institute of physics conference series, October 2012, pp. 73–94.
- [9] J. F. Donoghue, M. M. Ivanov, and A. Shkerin, *EPFL Lectures on General Relativity as a Quantum Field Theory*, ArXiv e-prints (February 2017), available at [arXiv:1702.00319](https://arxiv.org/abs/1702.00319).
- [10] L.D. Faddeev and V.N. Popov, *Feynman diagrams for the yang-mills field*, Physics Letters B **25** (1967), no. 1, 29–30.
- [11] F D M Haldane, *'luttinger liquid theory' of one-dimensional quantum fluids. i. properties of the luttinger model and their extension to the general 1d interacting spinless fermi gas*, Journal of Physics C: Solid State Physics **14** (1981), no. 19, 2585.
- [12] M. Henneaux and C. Teitelboim, *Quantization of gauge systems*, Princeton paperbacks, Princeton University Press, 1994.
- [13] G.'t Hooft, *A planar diagram theory for strong interactions*, Nuclear Physics B **72** (1974), no. 3, 461–473.
- [14] H. Kleinert, *Path integrals in quantum mechanics, statistics, polymer physics, and financial markets*, EBL-Schweitzer, World Scientific, 2009.
- [15] Z. Komargodski and A. Schwimmer, *On renormalization group flows in four dimensions*, Journal of High Energy Physics **12** (December 2011), 99, available at [arXiv:1107.3987](https://arxiv.org/abs/1107.3987).
- [16] A. Luther and I. Peschel, *Calculation of critical exponents in two dimensions from quantum field theory in one dimension*, Phys. Rev. B **12** (1975Nov), 3908–3917.
- [17] J. Maldacena, *The Large- N Limit of Superconformal Field Theories and Supergravity*, International Journal of Theoretical Physics **38** (1999), 1113–1133, available at [arXiv:hep-th/9711200](https://arxiv.org/abs/hep-th/9711200).
- [18] S. Mandelstam, *Soliton operators for the quantized sine-gordon equation*, Phys. Rev. D **11** (1975May), 3026–3030.
- [19] Daniel C. Mattis and Elliott H. Lieb, *Exact solution of a manyfermion system and its associated boson field*, Journal of Mathematical Physics **6** (1965), no. 2, 304–312, available at [arXiv:https://doi.org/10.1063/1.1704281](https://doi.org/10.1063/1.1704281).
- [20] J. Polchinski, *String theory: Volume 1, an introduction to the bosonic string*, Cambridge Monographs on Mathematical Physics, Cambridge University Press, 2005.
- [21] N. Seiberg and E. Witten, *Electric-magnetic duality, monopole condensation, and confinement in $N=2$ supersymmetric Yang-Mills theory*, Nuclear Physics B **426** (September 1994), 19–52, available at [arXiv:hep-th/9407087](https://arxiv.org/abs/hep-th/9407087).
- [22] H. E. Stanley, *Dependence of critical properties on dimensionality of spins*, Phys. Rev. Lett. **20** (1968Mar), 589–592.
- [23] S. Weinberg, *The quantum theory of fields*, The Quantum Theory of Fields 3 Volume Hardback Set, Cambridge University Press, 1995.

- [24] J. Wess and J. Bagger, *Supersymmetry and supergravity*, Princeton series in physics, Princeton University Press, 1992.
- [25] M. Wise and V. Bhansali, *Fixed income finance: A quantitative approach*, McGraw-Hill Education, 2010.
- [26] Edward Witten, *Quantum field theory and the jones polynomial*, Communications in Mathematical Physics **121** (1989Sep), no. 3, 351–399.
- [27] EDWARD WITTEN, *Introduction to cohomological field theories*, International Journal of Modern Physics A **06** (1991), no. 16, 2775–2792, available at [arXiv:http://www.worldscientific.com/doi/pdf/10.1142/S0217751X91001350](http://www.worldscientific.com/doi/pdf/10.1142/S0217751X91001350).
- [28] J. Zinn-Justin, *Renormalization of Gauge Theories and Master Equation*, Modern Physics Letters A **14** (1999), 1227–1235, available at [arXiv:hep-th/9906115](http://arxiv.org/abs/hep-th/9906115).
- [29] Jean Zinn-Justin, *Quantum field theory and critical phenomena*, Int. Ser. Monogr. Phys. **77** (1989), 1–914.