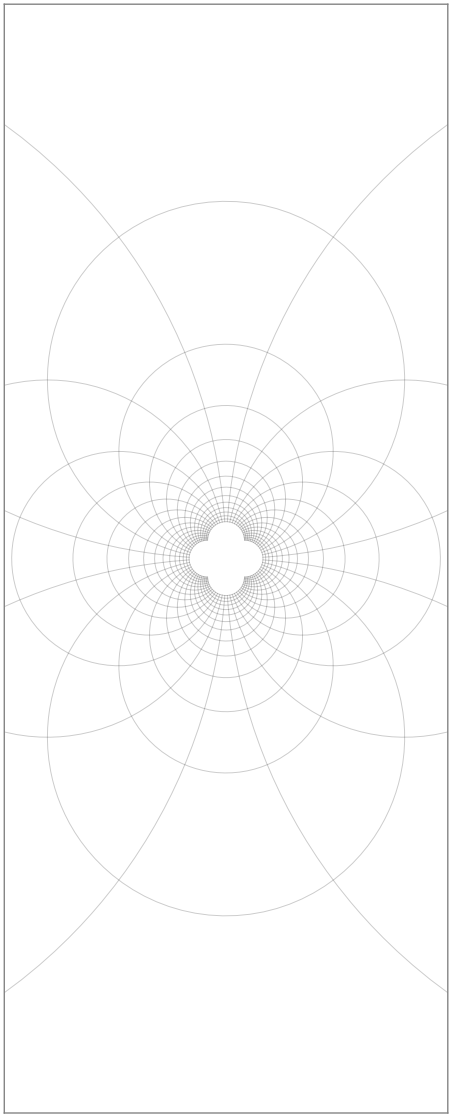
Log 2024

Lucas Z. Brito
April 23, 2024



Contents

Chapter 1

January

Contents

1.1 An Explicit Abelianization

1/9/24

Dihedral group The simplest non-abelian group available is the dihedral group $\mathrm{D}_6 \simeq \mathbb{Z}_2 \ltimes \mathbb{Z}_3$, the symmetry group of a triangle. The $\mathbb{Z}_2$ consists of the reflection of the triangle, and the $\mathbb{Z}_3$ is the rotations by $2\pi/3$ degrees. Thus we have

$$\mathrm{D}_6 \simeq \left\{e, r, r^2\right\} \ltimes \{e, p\}$$

It's not too hard to draw a bunch of triangles and find that this group has the following Cayley table:

	r	r^2	p	rp	r^2p
r	r^2	e	rp	r^2p	p
r^2	e	r	r^2p	p	rp
p	r^2p	rp	e	r^2	r
rp	p	r^2p	r	e	r^2
r^2p	rp	p	r^2	r	e

Here the nontrivial identities which require some toying around with the triangles are $pr = r^2p$, $pr^2p = r$, $pr^2 = rp$.

The semidirect What is the semidirect product structure on $\mathbb{Z}_2$ and $\mathbb{Z}_3$ that produces D_6? We can do this with a procedure identical to that which we performed on the $\mathrm{U}(1) \ltimes \mathbb{Z}_2$ symmetry of the particle on a ring of the 2023 logs. Now, there are two identities that capture the non-commutativity of the two symmetries, $pr^2p = r$ and $prp = r^2$ (two comments: first, notice this is like $PT_\alpha P = T_{-\alpha}$ for $\mathrm{U}(1) \ltimes \mathbb{Z}_2$, and also notice that these are precisely the

commutators with $p^{-1} = p$). We find

$$pr^2p = (p,e)\cdot(e,r^2)\cdot(p,e) = (p,e)\cdot(p,r^2\varphi_e(e)) = (p^2,\varphi_p(r^2)) = (e,r)$$
$$prp = (p,e)\cdot(p,\varphi_p(r)) = (p^2,r^2) = (e,r^2)$$

provided $\varphi_p(r^2) = r$—we don't even have to require $\varphi_p(r) = r^2$ because there is no other choice in this automorphism of $\mathbb{Z}_3$.

Commutator The commutator subgroup $[G,G]$ is the subgroup generated by the commutators of G. If the group is abelian, the commutator subgroup is trivial. The commutator is $[a,b] = a^{-1}b^{-1}ab$; indeed it measures how the conjugation by a changes b (or vice-versa): if a and b commute

$$[a,b] = e \quad \Longleftrightarrow \quad a^{-1}ba = b.$$

Abelianization Let's compute the abelianization $\mathrm{D}_6/[\mathrm{D}_6,\mathrm{D}_6]$ by finding what $[\mathrm{D}_6,\mathrm{D}_6]$ is. We have

$$[r,r^2] = e, \qquad [p,r] = prp^{-1} = prpr^2 = r^2r^2 = r,$$
$$[r,p] = rpr^2p = r^2, \qquad [p,r^2] = pr^2pr = r^2.$$

Evidently r generates the commutator subgroup and we have that $[\mathrm{D}_6,\mathrm{D}_6] \simeq \mathbb{Z}_3$ and thus the abelianization is

$$\mathrm{D}_6/[\mathrm{D}_6,\mathrm{D}_6] \simeq \mathbb{Z}_2,$$

which makes sense because when the group fails to commute the difference in the resulting actions is a rotation.

1.2 Central Charge and Second Cohomology 1/15/24

Blumenhagen and Plauschinn:

> Is is generally true that $H^2(\mathfrak{g},\mathbb{C})$ classifies the central extensions of an algebra $\mathfrak{g}$ modulo redefinitions of the generators. However, for semi-simple finite dimensional Lie algebras, one finds that their second cohomology group vanishes and so in this case there do not exist any central extensions.

1.2.1 Group Cohomology

Intuitively, group (co)homology can be be understood as the (co)homology of the classifying space of the group G, BG, which is the space whose fundamental group is G: $\pi_1(\mathrm{B}G) = G$ (the Eilenberg-MacLane theorem guarantees us such a space). For the purposes of computation,

however, we need a more concrete construction akin to a simplicial complex structure on a topological space.

TODO classifying space vs eilenberg maclane space?

Simplicial resolution The so-called **simplicial resolution**[1] does the job in a fairly intuitive way: we think of elements of G as vertices in a simplex, with n-simplices formed of n-tuples of elements of G, and consider the simplicial chain complex of this simplex. To be specific,

$$C_n = \bigoplus_{G^{n+1}} \mathbb{Z}$$

whose basis elements are $(g_0, \dots, g_n) \in G^n$, and with boundary map identical to that of simplicial complexes:

$$\begin{aligned} \partial_n C_n &\longrightarrow C_{n-1} \\ (g_0, \dots, g_n) &\longmapsto \sum_{i=0}^{n} (-1)^i (g_0, \dots, \hat{g}_i, \dots, g_n) \end{aligned}$$

Bar resolution An alternative resolution (which will come into play shortly) is the **bar resolution**

$$\bar{C}_n = \bigoplus_{G^n} \mathbb{Z}G$$

(the practical difference between this and the above direct sum is not manifest to me) whose basis elements are denoted $[g_1|g_2|\cdots|g_n]$. Here $\mathbb{Z}G$ is a $\mathbb{Z}G$-module—defined as an action of G on the abelian group $\mathbb{Z}$ satisfying $g(a+b) = ga + gb$, we can think of elements of $\mathbb{Z}G$-modules as simplicial chains in the sense that

$$g_1 n_1 + g_2 n_2 + \cdots + g_N n_N, \qquad n_i \in \mathbb{Z}$$

is in a $\mathbb{Z}G$ module. This comes with its own boundary map

$$\begin{aligned} \bar{\partial}_n \bar{C}_n &\longrightarrow \bar{C}_{n-1} \\ [g_1 \mid \cdots \mid g_n] &\longmapsto g_1[g_2 \mid \cdots \mid g_n] \\ &\quad + \sum_{i=1}^{n-1} (-1)^i [g_1 \mid \cdots \mid g_{i-1} \mid g_i g_{i+1} \mid g_{i+1} \mid \cdots \mid g_n] \\ &\quad + (-1)^n [g_1 \mid \cdots \mid g_{n-1}] \end{aligned}$$

Equivalence of resolutions (It is the corresponding coboundary map of the bar boundary map that will appear in the classification of the central extension). The two constructions can be related by a map

$$[g_1 \mid \cdots \mid g_n] \mapsto (e, g_1, g_1 g_2, \dots, g_1 g_2 \cdots g_n)$$

[1] For the present purposes, it suffices to think of a resolution as a chain complex; indeed it has the same structure as an exact sequence.

which actually is an isomorphism, meaning that the bar and simplicial resolutions are entirely isomorphic and their (co)homologies will agree. Actually, this is an instance of a broader fact—these resolutions are both **projective**, and all projective resolutions are chain homotopy equivalent $\implies$ their (co)homologies agree.

1.2.2 Central extensions from cohomology

Consider the central extension $\tilde{G}$ of G by Z:

$$0 \longrightarrow Z \longrightarrow \tilde{G} \longrightarrow G \longrightarrow 0$$

Now, to differentiate the abelian group composition of Z from the nonabelian composition of G and $\tilde{G}$, we will write Z as the formal exponentiation of some additive abelian group $Z = \exp(F)$. (This is also in anticipation of the interpretation of Z as a phase in some field F, e.g., $\mathbb{R}$.)

We are guaranteed that $G \simeq \tilde{G}/\exp(F)$, although this isomorphism is not canonical—one way you can think of this is we need to choose which phase will go on each element of G. To be formal, $\tilde{G}/\exp(F) = \{\tilde{g}\exp(F) \mid \tilde{g} \in \tilde{G}\}$; the isomorphism assigns each element of g to a representative of a coset $\tilde{g}\exp(F)$ (i.e., an arbitrary element of that coset, meaning an element of $\tilde{g}$ with some phase); such a set—one composed of a representative from each coset—is termed a *section*.

At any rate, having made such a choice, the isomorphism is canonical. Let it be $\gamma : G \to \tilde{G}/Z$. Then every element of $\tilde{G}$ can be written as some $\gamma(g)\exp(a)$, $g \in G$, $a \in F$, and the composition is generally

$$\gamma(g_1)\gamma(g_2) = \gamma(g_1 g_2)\exp(c(g_1, g_2)) \tag{1.1}$$

with c some function $G \times G \to F$. Now, generally functions $G^n \to F$ are group n-cochains (strictly speaking F needs to be a G-module, though I'm not sure how G acts on F in this case), so c is a group 2-cochains, but not one that is uniquely determined by $\tilde{G}$. Specifically, we can choose another section, say γ', under which

$$\gamma'(g_1)\gamma'(g_2) = \gamma'(g_1 g_2)\exp(c'(g_1, g_2)) \tag{1.2}$$

Now, in transforming to our original section γ the worst that can happen is we acquire a phase:

$$\gamma'(g) = \gamma(g)\exp(d(g))$$

for some $d : G \to F$. Then we have eq. (1.2) becomes

$$\begin{aligned}\gamma(g_1)\exp(d(g_1))\gamma(g_2)\exp(d(g_2)) &= \gamma(g_1 g_2)\exp(d(g_1 g_2))\exp(c'(g_1, g_2)) \\ \gamma(g_1)\gamma(g_2) &= \gamma(g_1 g_2)\exp\left[d(g_1 g_2) - d(g_1) - d(g_1) + c'(g_1, g_2)\right]\end{aligned}$$

Comparing with eq. (1.1) we find

$$c(g_1, g_2) - c'(g_1, g_2) = d(b_1 b_2) - d(b_1) - d(b_2)$$

The right hand side is nothing but the (bar) coboundary map $\delta d(b_1, b_2)$, so we have found that that any two central extensions are determined only up to a coboundary.

On the other hand, consider the following composition:

$$\begin{aligned}\gamma(g_1)\gamma(g_2)\gamma(g_3) &= \gamma(g_1 g_2)\gamma(g_3)\exp(c(g_1, g_2)) \\ &= \gamma(g_1 g_2 g_3)\exp(c(g_1, g_2))\exp(c(g_1 g_2, g_3))\end{aligned}$$

But associativity tells us

$$\begin{aligned}\gamma(g_1)\gamma(g_2)\gamma(g_3) &= \gamma(g_1)\gamma(g_2 g_3)\exp(c(g_1, g_2)) \\ &= \gamma(g_1 g_2 g_3)\exp(c(g_2, g_3))\exp(c(g_1, g_2 g_3))\end{aligned}$$

Equating the two right hand sides:

$$\begin{aligned}&\gamma(g_1 g_2 g_3 \exp(c(g_1, g_2))\exp(c(g_1 g_2, g_3))) = \gamma(g_1 g_2 g_3)\exp(c(g_2, g_3))\exp(c(g_1, g_2 g_3)) \\ &\quad\implies \exp(c(g_1, g_2) + c(g_1 g_2, g_3) - c(g_2, g_3) - c(g_1, g_2 g_3)) = 1 \\ &\quad\implies c(g_1, g_2) + c(g_1 g_2, g_3) - c(g_2, g_3) - c(g_1, g_2 g_3) = 0\end{aligned}$$

again, this is in fact nothing but the coboundary map $\delta c(g_1, g_2, g_3)$, so that c is in fact a cocycle (it has vanishing coboundary).

Cohomology What we've learned is that the central extension is determined by a 2-cocycle c up to the addition of a 2-coboundary d, thus it is classified by the second group cohomology $H^2(G, F) = \ker\delta/\mathrm{im}\,\delta$. Now, the remarkable fact is that this is also the case at the level of the Lie algebra—the central extensions (properly, the central charge) is classified by the second Lie algebra cohomology $H^2(\mathfrak{g})$. We can also understand this by noticing that the central charge is nothing more than the generator of Z—it is the generator of the central extension, hence the terminology "central charge" (it is the generator—the charge—of the center). It is a consequence of the Whitehead lemma that $H^2(\mathfrak{g}) = 0$ if $\mathfrak{g}$ is semisimple, meaning the semisimple Lie algebras cannot carry a central charge.

Here is a quick illustration of how the above considerations play with central charges as they appear at the level of the algbera. Provided $[X, Y]$ is central,

$$\begin{aligned}e^X e^Y e^{-X} e^{-Y} &= e^{X+Y+\frac{1}{2}[X,Y]} e^{-X} e^{-Y} \\ &= e^{X+Y+\frac{1}{2}[X,Y]-X+\frac{1}{2}[X,Y+\frac{1}{2}[X,Y],-X]} e^{-Y} \\ &= e^{Y+\frac{1}{2}[X,Y]-\frac{1}{2}[Y,X]} e^{-Y} \\ &= e^{Y+[X,Y]-Y+\cancel{\frac{1}{2}[[X,Y],Y]}} \\ &= e^{[X,Y]}\end{aligned}$$

(this is a neat formula relating the algebra commutator to the group commutator). Then

$$\begin{aligned} e^{\gamma([X,Y])} &= \gamma(e^X)\gamma(e^Y)\gamma(e^{-X})\gamma(e^{-Y}) \\ &= \gamma(e^X e^Y e^{-X} e^{-Y}) e^{c(e^X, e^Y)} e^{c(e^X e^Y, e^{-X})} e^{c(e^X e^Y e^{-X}, e^{-Y})} \\ &= \gamma(e^{[X,Y]}) e^{C(X,Y)} \end{aligned}$$

Then

$$\gamma([X,Y]) = [X,Y] + C(X,Y)$$

References

- Clara Löh, Group Cohomology.
- user568976, Understanding the relation between central extensions and 2-cocycles and coboundaries, URL (version: 2020-04-17): https://math.stackexchange.com/q/3629722

Chapter 2

February

Contents

2.1 QFTs from Neural Networks

2/1/24

In QFT topological defects are fixed points of RG flow—if learning is like an rg flow, what happens if there's a topological defect in a NN?

2.2 Chern-Simons Theory

2/5/24

Here's a silly mnemonic to remember the Chern-Simons term: the godmother of computing, Ada Lovelace, was inspired by the mechanism of mechanical looms, which performed weaves and braids not unlike those of the anyons that populate a Chern-Simons theory. The term $A \wedge \mathrm{d}A$, is associated as $A \wedge \mathrm{d}A \to \mathrm{AdA} \to$ Ada Lovelace.

(What's that? Doesn't the wedge product of any form with itself vanish? No, silly, recall these are *Lie algebra-valued one forms*, for which the wedge product is essentially the

commutator. If this theory was Abelian, however, the wedges *would* vanish.)

2.2.1 Anyons in the semiclassical theory

$$\mathcal{L} = \frac{k}{4\pi}\epsilon^{\mu\nu\rho}A_\mu\partial_\nu A_\rho + A_\mu J^\mu$$

$$\frac{\delta S}{\delta A_\mu} = 0 \Longrightarrow 0 = \frac{k}{4\pi}\epsilon^{\mu\nu\rho}\partial_\nu A_\rho + J^\mu \Longrightarrow \frac{k}{4\pi}\epsilon^{\mu\nu\rho}\partial_\nu A_\rho = -J^\mu$$

$$\frac{k}{4\pi}\epsilon^{ij}\partial_i A_j = J^0 = \rho \Longrightarrow \frac{k}{2\pi}B = \rho$$

$$\frac{k}{4\pi}\epsilon^{ij}(\partial_j A_0 - \partial_0 A_j) = -J^i \Longrightarrow \frac{k}{4\pi}\epsilon^{ij}E_j = J^i$$

The Wilson lines obey the algebra

$$W_x W_y = e^{2\pi i/k} W_y W_x$$

But note that this implies W_y^k commutes with W_x:

$$W_x W_y^k = (e^{2\pi i/k})^k W_y^k W_x \Longrightarrow W_x W_y^k = W_y^k W_x$$

and, of course, W_y^k commutes with any power of W_y, so W_y^k is in the center, which by Schur's lemma implies that it is proportional to the identity. It is the Casimir of this algebra

Now consider some state $|n\rangle$ in this representation labelled by quantum number n so that $W_x|n\rangle = \lambda_n|n\rangle$. Then

$$W_x W_y|n\rangle = e^{2\pi i/k} W_y W_x|n\rangle = e^{2\pi i/k} W_y \lambda_n|n\rangle = e^{2\pi i/k}\lambda_n W_y|n\rangle$$

So $W_y|n\rangle$ has eigenvalue $e^{2\pi i/k}\lambda_n|n\rangle$ and W_y acts as a raising operator. But $W_y^k = 1$, so one can only move up in the weight lattice N times. Thus we have a k-dimensional irrep, and a k-degenerate groundstate. In fact, this is the irrep of the 1-form symmetry generated by the Wilson lines CS theory enjoys.

2.2.2 Large-k partition function and framing anomaly

We study the nonabelian Chern-Simons partition function at large k, following Witten (1989) and Witten and Bar-Natan (1991). Analogous to the large-N approximation, large-k Chern-Simons theory justifies approximating the partition function by its saddle-point—i.e., its classical equation of motion. Since for Chern-Simons theory this equation of motion is $F^{\mu\nu} = 0$, the saddle-point corresponds to the "*moduli space*" of flat connections.[1]

$$(A^{(\alpha)} + a) \wedge \mathrm{d}(A^{(\alpha)} + a) = A^{(\alpha)} \wedge \mathrm{d}A^{(\alpha)} + A^{(\alpha)} \wedge \mathrm{d}a + a \wedge \mathrm{d}A^{(\alpha)} + a \wedge \mathrm{d}a$$

[1] Physicists can think of moduli spaces as the space of configurations satisfying a certain requirement.

(temporarily suppressing the (α) index)

$$\begin{aligned}(A+a)\wedge(A+a)\wedge(A+a) &= (A+a)\wedge[A\wedge A + A\wedge a + a\wedge A + a\wedge a]\\ &= A\wedge A\wedge A + A\wedge A\wedge a + A\wedge a\wedge A + A\wedge a\wedge a\\ &\quad + a\wedge A\wedge A + a\wedge A\wedge a + a\wedge a\wedge A + a\wedge a\wedge a\end{aligned}$$

Treating a as small, we ignore the third-order term

$$\begin{aligned}&= A\wedge A\wedge A + \cancel{A\wedge A\wedge a} + A\wedge a\wedge A + \cancel{A\wedge a\wedge a}\\ &\quad + \cancel{a\wedge A\wedge A} + a\wedge A\wedge a + \cancel{a\wedge a\wedge A}\end{aligned}$$

The indicated terms cancel because for the wedge product of Lie algebra-valued forms we have

$$\alpha\wedge\beta = (-1)^{1+\deg\alpha\cdot\deg\beta}\beta\wedge\alpha$$

and recall that since we are expanding about the saddle-point, linear fluctuations vanish and we have

$$= A\wedge A\wedge A + a\wedge A\wedge a$$

Thus we are left with

$$\begin{aligned}\mathcal{L} &= A^{(\alpha)}\wedge \mathrm{d}A^{(\alpha)} + A^{(\alpha)}\wedge A^{(\alpha)}\wedge A^{(\alpha)} + a\wedge(\mathrm{d}a + A^{(\alpha)}\wedge a)\\ &= \mathcal{L}_{\mathrm{CS}}[A^{(\alpha)}] + a\wedge Da = \mathcal{L}_{\mathrm{CS}}[A^{(\alpha)}] + \epsilon^{\mu\nu\rho}a_\mu D_\nu a_\rho\end{aligned}$$

where we've identified a covariant derivative for the saddle point field configuration $D = \mathrm{d} + A\wedge\cdot = \partial_\mu + [A_\mu,\cdot]$.

Now, we introduce the *elliptical operator* $\star D + D\star$. This operator manifestly maps odd-degree forms to odd-degree forms and even-degree forms to even-degree forms, so one can consider its action on the space of odd forms strictly, which we call L_-. In three-dimensions L_- acts on $\Omega^1(M)\oplus\Omega^3(M)$; we thus combine (a,ϕ) into an element of $\Omega^1(M)\oplus\Omega^3(M)$ and observe that then

$$\begin{aligned}(a,\phi)\wedge\star(\star D + D\star)(a,\phi) &= (a,\phi)\wedge\star(\star Da + D\star\phi, \cancel{\star D\phi} + D\star a)\\ &= a\wedge Da + \cancel{a\wedge\star D\star\phi} + \phi\star D\star a\end{aligned}$$

(Notice that L_-'s action is a bit subtle here—we get mixing of components in the sense that $a\in\Omega^1$ but $D\star a\in\Omega^3$, so we must take it to the second component.) The indicated term vanishes in the last line because ϕ is non-dynamical. Also, if $\phi\star D\star a$ looks suspicious, notice that $\star D\star a$ is a zero form, i.e., a scalar, so this is just scalar multiplication on ϕ.

2.2.3 1-form symmetry

2.2.4 Spin manifolds

Consider integration by parts for forms:

$$\int_X d\alpha \wedge \beta = \int_X \left[d(\alpha \wedge \beta) - (-1)^{\deg \alpha} \alpha \wedge d\beta \right] = \int_{\partial X} \alpha \wedge \beta - (-1)^{\deg \alpha} \int_X \alpha \wedge d\beta$$

Thus setting $\alpha = A$ and $\beta = dA$, one finds

$$\int_X F \wedge F = \int_{\partial X} A \wedge dA - (-1)^{\deg \alpha} \int_X \cancel{A \wedge d^2 A}$$

The argument is that for this extension to be consistent, the partition function needs to be the same for any two X_1 and X_2 sharing M as a boundary. Thus, we may glue two such X_1 and X_2 along M (so that the integral becomes an integration over Y the manifold resulting from the gluing):

$$\Delta S_{\text{CS}} = S_{\text{CS}}^{X_1} - S_{\text{CS}}^{X_1} = \frac{k}{2\pi} \int_Y F \wedge F$$

If we stipulate

$$\frac{k}{4\pi} \int_Y F \wedge F \in 2\pi\mathbb{Z} \iff \frac{k}{8\pi^2} \int_Y F \wedge F \in \mathbb{Z}$$

we have $e^{\Delta S_{\text{CS}}} = e^{i2\pi n} = 1$ and the partition function is invariant. When is the above true? As it turns out, exactly when Y is a spin manifold, which leads to the integral resolving to an integer

$$\int_Y F \wedge F \in 2\mathbb{Z} \implies \Delta S_{\text{CS}} = 2\pi k \cdot n \in 2\pi\mathbb{Z}$$

When Y does not meet this condition, the integral is simply an integer

$$\int_T F \wedge F \in \mathbb{Z} \implies \Delta S_{\text{CS}} = 2\pi k \cdot \frac{n}{2} \in 2\pi\mathbb{Z} \text{ only if } k \in 2\mathbb{Z}$$

i.e., the Chern-Simons level must be *even* in addition to being an integer.

How could this possibly be true? First things first: here we can view $F/2\pi$ as a member of $H^2(Y;\mathbb{Z})$. $dF = 0$ by the Levi-Civita identity—i.e., because $F = dA$—so it is in the kernel of the coboundary map, and in this context

$$\begin{aligned}(F + d\alpha) \wedge (F + d\alpha) &= F \wedge F + d\alpha \wedge F + F \wedge d\alpha + d\alpha \wedge d\alpha \\ &= F \wedge F + d^2\alpha \wedge \alpha + \underbrace{\cancel{2d(\alpha \wedge F)} + \cancel{d(\alpha \wedge d\alpha)}}_{\text{no } \partial X}\end{aligned}$$

so it is only defined up to an exact form. Lastly, $F/2\pi \in H^2(Y, \mathbb{Z})$ because the integral of F is valued in $2\pi\mathbb{Z}$ thanks to the Dirac quantization condition. Recall that in this case elements of the cohomology implicitly come up with an integration—this is the appropriate approach for differential cohomology, for instance De Rham cohomology. Correspondingly, cup products turn into wedge products.

We can make contact with the Stiefel-Whitney class, we consider the reduction

$$H^2(Y; \mathbb{Z}) \to H^2(Y; \mathbb{Z}_2)$$

so the lift of the Stiefel-Whitney class w_2 is only defined mod 2. For instance w_2 vanishing corresponds to it being even valued in $H^2(Y; \mathbb{Z})$. Additionally, the intersection form satisfies

$$\frac{F}{2\pi} \smile \frac{F}{2\pi} = w_2 \smile \frac{F}{2\pi}$$

(TODO why is this true?). Since $F/2\pi \in \mathbb{Z}$ and $w_2 \in 2\mathbb{Z}$ if the manifold is spin, one has the intersection form is an even integer, and

$$\int_Y \frac{F}{2\pi} \wedge \frac{F}{2\pi} \in 2\mathbb{Z}$$

and when the manifold is not spin, it is simply integer-valued.

References

- Sachdev, "Topology and Chern-Simons Theories," PHYS268.
- Witten, "Perturbative expansion of Chern-Simons theory with noncompact gauge group."
- Witten, "Quantum Field Theory and the Jones Polynomial."
- Ethan Lake, *QFT Diary*.
- Santiago Quintero de los Ríos "Seiberg-Witten Theory and the Intersection Form"

2.3 Spin Manifold Miscellany 2/11/24

2.3.1 Vierbeins, Tetrads, Spin Connections

The vierbein (German for "four-legged") formalism is an alternative formulation of a coordinate system on a curved manifold which is intimately related to the spin connection. They are defined as the eigenvalues of the metric:

$$g_{\mu\nu} = e^a_\mu(x)\eta_{ab}e^b_\nu(x)$$

Notice that, viewing the metric as a background field, the vierbein formalism essentially defines a field $e^a_\mu(x)$, which we interpret as the (local) coefficients for a vector expressed in a *fixed* orthonormal basis (referred to as the **tetrad** basis).

Now, the more formal way to define the spin connection is as follows: There is a short exact sequence

$$0 \longrightarrow \mathbb{Z}_2 \longrightarrow \mathrm{Spin}(N) \xrightarrow{\phi} \mathrm{SO}(N) \longrightarrow 0^2 \tag{2.1}$$

which restates the familiar mantra "spinors are the square root of vectors." Now, we might consider a $\mathrm{Spin}(N)$ spin bundle related to an $\mathrm{SO}(N)$ vector bundle by ϕ. Where the Levi-Civita connection is the connection appropriate for vectors—i.e., it can be thought of as an $\mathrm{SO}(N)$ action—the spin-connection is the connection on spinors $\in \mathrm{Spin}(N)$ induced by the pullback of ϕ. To be sure that everything is well defined, there is a Levi-Civita one form

$$\omega_{ij}(X) = g(\nabla_X e_i, e_j) \qquad \text{(corresponding to } D_\mu\text{)}$$

the pullback $\phi^*\omega$ of which is the spin connection. In this construction N is the dimension of the manifold (since $\mathrm{SO}(N)$) is the structure group of the orthonormal frame bundle, which seems to place some constraints on the spinor bundle $\mathrm{Spin}(N)$ that you can put on the manifold (though maybe you can fiber $N > \dim M$ bundles on it through an immersion?).

Another way to think of why we might need the spin connection: as it turns out, there are no finite-dimensional spinor representations of the general covariance, or autoisometry group $\mathrm{GL}_4(\mathbb{R})$ (LOOK INTO THIS) that gravity enjoys as a symmetry. (Some authors refer to this as the diffeomorphism group, but they are WRONG! See section 2.4.) Since we know that the Lorentz group admits spinor representations, the trick is to extract the curvature of the manifold into the tetrad basis $e^\mu_a(x)$, leaving behind the "internal" flat metric η_{ab} which of course transforms under the Lorentz group, thus allowing us to define spinorial transformations in a curved spacetime according to the covariant derivative:

$$D_\mu = \partial_\mu - \frac{i}{4}\omega^{ab}_\mu \sigma_{ab}, \qquad \sigma_{ab} = \frac{i}{2}\left[\gamma^a, \gamma^b\right]$$

and the spin connection is

$$\omega^a_{\mu b} = e^a_\nu e^\lambda_b \Gamma^\nu_{\mu\lambda} - e^\lambda_b \partial_\mu e^a_\lambda.$$

This is the covariant derivative that would appear in, e.g., the Dirac equation on a curved background.

2.3.2 Spinor Reps of GL_4

TODO

[2] This is silly, but for a long time I wondered why we used $\mathrm{Spin}(N)$ instead of $\mathrm{SU}(N)$ for spinors. Well, this only holds for $N = 3$, where $\mathrm{Spin}(3) \simeq \mathrm{SU}(2)$. On the other hand $\mathrm{Spin}(4) \simeq \mathrm{SU}(2) \times \mathrm{SU}(2)$, hence the two Weyl spinors in relativistic theories.

2.3.3 Stiefel-Whitney classes

Characteristic classes measure how bundles over a manifold twist, i.e., they measure obstructions to a global trivialization. For a manifold M, the class c associates an element of the cohomology $c(P) \in H^k(M)$ to a given G-bundle P (properly we should consider P a representative of an isomorphism class of G-bundles). The characteristic class that measures obstructions to our ability to define global orthonormal frame over the manifold is the first Stiefel-Whitney class. Put it simpler terms, the first Stiefel-Whitney class measures the failure of the manifold to be orientable.

Analogously, the second Stiefel-Whitney class measures the failure of a manifold to admit a spin structure, that is, a manifold is a spin manifold iff its second Stiefel-Whitney class is trivial. We will see how much like how $\mathrm{Spin}(M)$ is the double cover of $\mathrm{SO}(M)$, the second Stiefel-Whitney class is sort of the double cover of the first.

1st SW class The first Stiefel-Whitney class is defined as the determinant of the transition maps on the charts of the manifold. Consider two charts U_i and U_j related by the transition map t_{ij}, itself a member of the structure group $\mathrm{SO}(m = \dim M)$. $(i, j$ here do not index the entries of the matrix but instead label the charts.) We define a 1-cochain with coefficients in $\mathbb{Z}_2$ by the determinant of the transition function:

$$f(i,j) \equiv \det(t_{ij}) = \pm 1 \qquad f \in C^1(M, \mathbb{Z}_2)$$

(This is a Čech cocycle). Now, $t_{ij}t_{jk}t_{ji} = 1$ because this transitions us back to the first chart i. Consequently the coboundary map sends

$$(\delta f)(i,j,k) = \det t_{ij} \det t_{jk} \det t_{ki} = \det(t_{ij}t_{jk}t_{ki}) = 1$$

And, further, if we consider transforming to a different frame on U_i by $h_i \in \mathrm{O}(m)$—and likewise for U_j—the transition functions transform to $t_{ij} \to h_i t_{ij} h_j^{-1}$, in which case we find

$$f(i,j) \to \det(h_i t_{ij} h_j^{-1}) = \underbrace{\det h_i \det h_j}_{\equiv(\delta f_0)(i,j)} \det f_{ij}$$

Here we've defined $f_0(i) = \det h_i$ to be a 0-cochain. Thus f transforms by an exact 1-chain (which vanishes upon taking the coboundary map), and the equivalence class $[f]$ of f up to a 0-coboundary is evidently a member of $H^1(M; \mathbb{Z}_2)$. This is the Stiefel-Whitney class, $w_1(M) = [f]$.

1st SW and Orientability Now, the first Stiefel-Whitney class is trivial—all functions in $[f]$ are identically 1—iff the manifold is orientable. If M is orientable the structure group is $\mathrm{SO}(m)$ and $\det(t_{ij})$ is necessarily unity. Meanwhile, if $w_1(M) = 1$, $f(i,j) = (\delta f_0)(i,j)$ and we can always choose h_i and h_j such that this product is unity (i.e., the transition map continues to posses unit determinant).

2nd SW class The second Stiefel-Whitney class measures a slightly different obstruction: instead of taking the consistency of the transition maps for granted (and thus taking the existentence of a tangent bundle for granted) and measuring their failure to preserve orientation as the first class does, the second measures the failure of the transition maps to be consistent in the first place. This is the sense in which it tells us whether we can put a consistent spin structure on the manifold.

Consider the lift of $t_{ij} \in \mathrm{SO}(m)$ to the double cover $\tilde{t}_{ij} \in \mathrm{Spin}(m)$, i.e., $\tilde{t}_{ij}$ such that $\phi(\tilde{t}_i j) = t_{ij}$ for the two-to-one homomorphism given in eq. (2.1). We of course need

$$\tilde{t}_{ij}\tilde{t}_{jk}\tilde{t}_{ki} = I$$

for the spin bundle to be well-defined. We naturally also have

$$\phi(\tilde{t}_{ij}\tilde{t}_{jk}\tilde{t}_{ki}) = t_{ij}t_{jk}t_{ki} = I$$

Given the above and eq. (2.1), $\tilde{t}_{ij}\tilde{t}_{jk}\tilde{t}_{ki}$ can only fail to be the identity up to a multiplicative constant—in fact, it can only fail to be the identity up to a sign. Thus, we define a 3-cochain f satisfying

$$\tilde{t}_{ij}\tilde{t}_{jk}\tilde{t}_{ki} = f(i,j,k)I$$

Further, consider the two-cochain defined as the sign of the transition function $\tilde{t}_{ij}$. Consider picking an arbitrary primage of $\phi(t_{ij})$ that might have a sign in front of it. Then we generalize to

$$\tilde{t}_{ij}\tilde{t}_{jk}\tilde{t}_{ki} = f(i,j,k)I \to \underbrace{\mathrm{sign}(t_{ij})\,\mathrm{sign}(t_{jk})\,\mathrm{sign}(t_{ki})}_{(\delta f)(i,j,k)} f(i,j,k)I.$$

Thus f represents an equivalence class $[f] \in H^2(M;\mathbb{Z}_2)$.

> **2nd SW and Spin Structures** Of course if the manifold is spin, $\tilde{t}_{ij}\tilde{t}_{jk}\tilde{t}_{ki} = 1$, so w_2 is trivial. Converesely, if w_2 is trivial, it is defined up to a transformation
>
> $$\tilde{t}_{ij}\tilde{t}_{jk}\tilde{t}_{ki} \to f(i,j)f(j,k)f(k,i)\tilde{t}_{ij}\tilde{t}_{jk}\tilde{t}_{ki}$$
>
> But if f is trivial it is necessarily equal to $(\delta f_1)(i,j,k)$, so we have
>
> $$\tilde{t}_{ij}\tilde{t}_{jk}\tilde{t}_{ki} \to [f(i,j)f(j,k)f(k,i)]^2 I = I.$$

References

- Nakahara, *Geometry, Topology and Physics* (2003), Section 11.6.

$O(n) \subset O(n+1)$. relative homotopy $\pi_k(O(n+1), O(n)) \cong \pi_k(S^n)$. $\pi_k(A) \cong \pi_k(\tilde{A})$ (generally true?).

Conformal anomaly relation to central charge

2.4 The Diffeomorphism Kerfuffle 2/16/24

People will commonly refer to general relativity theory as "diffeomorphism invariant" (or sometimes more inocuously, "general covariant"), which they insist is euphemism for "invariant wrt coordinate transformations." The culprit which motivated this entry was the claim that Chern-Simons theory is general covariant, but the framing anomaly breaks this covariance. Now, this case certainly helps to split the two notions—while gravity might be general covariant, it is certainly not diffeomorphism invariant in the sense that Chern-Simons theory is (independent of the metric and thus completely bendable).

Ethan Lake comes to the rescue: "...diffeomorphisms do not care about distances... such a theory would therefore need to be topological." He accuses the physics community of being terribly imprecise with terminology, and rightfully so: what people refer to as *diffeomorphism invariance* in fact refers to *isometries*, transformations which preserve the metric (in a global sense that would also preserve its derivatives and thus the curvature) and thus the Einstein-Hilbert action; this is the *true* symmetry of GR. To be explicit, an isometry $f : (M, g) \to (N, h)$ has

$$g(V, W) = h(f_* V, f_* W) \iff g_{\mu\nu} V^\mu W^\nu = h_{\alpha\beta} \frac{\partial y^\alpha}{\partial x^\mu} \frac{\partial y^\beta}{\partial x^\nu} V^\mu W^\nu$$

(f_* being the pullback of f). This means the metric transforms as

$$h_{\alpha\beta}(y) = \left[\frac{\partial y^\alpha}{\partial x^\mu} \frac{\partial y^\beta}{\partial x^\nu} \right]_x g_{\mu\nu}(x)$$

Properly, in GR we are concerned with what Lake calls "autoisometries"—we are sending the manifold to itself. *This* is the coordinate transformation people confuse for a diffeomorphism. We aren't changing the manifold, but generally change the metric by virtue of changing the coordinates.

WRITE A BIT ON THE KILLING EQUATION

References

- Ethan Lake, Manifesto on spacetime symmetries, diffeomorphisms, and conformal transformations.

2.5 Virtual Serial Ports for Development 2/20/24

I've been working on some software that reads a stream of bytes from a serial port—namely, an FPGA taking input from four photon detectors. The FPGA device is obviously not

something I can take home; nonetheless, I want to be able to develop the software outside the lab. Naturally, what I need is a virtual serial port, which can be simulated with `socat`:

```
socat -d -d pty,raw,echo=0 pty,raw,echo=0
```

2.6 Anomaly: Fermion on a Ring 2/20/24

2.6.1 't Hooft Anomalies and Cohomology

Homotopy π_k fails excision and additivity Eilenberg-Steenrod axioms.

https://www.damtp.cam.ac.uk/user/tong/susy/susyqm2.pdf

- Why $T(m) = -m$. It's not that deep, just work it out in the action. □
- Explain how APS implies

$$\lim_{m\to\infty} \frac{Z[+m, g]}{Z[-m, g]} = \exp\left(i \int_X \mathrm{CS}_{\mathrm{grav}}\right)$$

  Seems like something technical. In general look to understand the CS gravity term, see if there's relation between the above and the fact that it is the appropriate counterterm in the framing anomaly.

- How $\rho(m) + \rho(-m) = 1$?
- Why $Z[m] = 1$ at long distances equiv to theory trivially gapped? My guess is that this amounts to ignoring spectra above a certain energy scale, in which case there's just one state in the partition function.
- Understand the whole "nontrivial physics on interfaces" bit.

2.7 Anomaly: Compact Boson ...

2.8 Landau Poles

"Power divergence is strongest at the cutoff, logarithmic divergence receives equal contributions at all scales" https://www.youtube.com/watch?v=RiMcxSz-PnI (13:52)

2.9 Zeta Function Regularization ...

GAMMAM MATRICES FLIP SIGN UNDER TIME REVERSAL

$[Y, K(A, n)] \simeq H^n(Y, A)$ ($K(A, n)$ is Eilenberg-Maclane space.)

Chapter 3

March

Contents

3.1 Homotopy vs. Homology

3/5/24

Relationship between π_k and H_k. H_1 is abelianization of π_1. In general the relationship is not so clear. E.g., $\pi_2(\mathbb{R}\mathrm{P}^2) = \mathbb{Z}$, $H_2(\mathbb{R}\mathrm{P}^2) = 0$. This is because $\mathbb{R}\mathrm{P}^2 = e^0 \cup e^1 \cup e^2$ and

$$\mathbb{Z}\left\{e^2\right\} \xrightarrow[\partial]{2} \mathbb{Z}\left\{e^1\right\} \xrightarrow[\partial]{0} \mathbb{Z}\left\{e^0\right\}$$

thus $H_2(\mathbb{R}\mathrm{P}^2) = 0$. (boundary of $e^2 = 2e^1$).

Another example: $X = \mathbb{C}\mathrm{P}^\infty$ has

$$H_k(X) = \begin{cases} \mathbb{Z} & k \text{ even} \\ 0 & \text{else} \end{cases}$$

$$\pi_k(X) = \begin{cases} \mathbb{Z} & k = 2 \\ 0 & \text{else} \end{cases}$$

1. Spaces that are weakly h.e. (in particular have the same π_k) have the same homology groups. Given $f : X \to Y$ whe, wlog $f : X \hookrightarrow Y$ (mapping cylinder). WTS $H_k(Y, X) \equiv 0$, Fix $[z] \in H_k(Y, X)$ $z = \sum a_i \sigma_i$ $a_i \in \mathbb{Z}$, z cycle: $\partial z \in C_{k-1}(X)$, $\sigma_i : \Delta^k \to Y$. Build complex (K, L) from $z \bigcup_i \Delta^k / \sim \partial\, K \subset \partial K$. (take two simplices and glue along common boundary). By construction have map $g : (K, L) \to (Y, X)$ $[w] \in H_k(K, L)$ s.t. $g^*(w) = [z]$.

$$\begin{array}{ccc} (K,L) & \longrightarrow & (Y,X) \\ & \searrow & \uparrow \\ & & (X,X) \end{array}$$

   Since $H_k(X, X) = 0$ this implies $[z] = 0$. WTS $(K, L) \to (Y, X)$ homotopic to $(K, L) \to (X, X)$. Use compression, $\pi(Y, X) = 0$ to inductively construct homotopy as in proof of Whitehead.

2. If X is n-connected ($\pi_k = 0$ for $k \leq n$) then also $\tilde{H}_k = 0$ for $k \leq n$. Thm (cellular approx of spaces). Any space Z is whe to a cellular complex, i.e.g,, $\exists X$ cell complex and $f : X \to Z$if Z is n-connected then can assume X has no cells of dimensoin $1, \ldots, n$. Proof: X n-connected. By cellular approx $\exists Z$ cell complex and whe $Z \to X$ and Z has no cells in dim $1, \ldots, n$. By 1. suffices to show $\tilde{H}_k(Z) = 0$, $k \leq n$.

$$Z\,\{n+1\text{-cells}\} \to 0 \to \cdots \to 0 \to \mathbb{Z}\,\{0\text{-cells}\}$$

   using cellular homology $\tilde{H}_k(Z) = 0$ $k \leq w$.

3. (Hurewicz) If X is n-connected $n \geq 2$ then $\pi_{n+1}(X) \simeq H_{n+1}(X)$.

E.g., $\mathbb{C}\mathrm{P}^\infty$ 1-connected, $\pi_2(\mathbb{C}\mathrm{P}^\infty) = H_2(\mathbb{C}\mathrm{P}^\infty)$.

3.2 assortment of spin chain facts ...

Shorter correlation length -¿ gapped LSM. translational invariant so(3) interaction spin-1/2 chain is gapless. lieb-robinson bound gapless — power law corr, gapped - exponential corr.

3.3 Fibrations 3/9/24

Homotopy lifting property A map between two spaces $p : E \to B$ is said to have the homotopy lifting property for a space X, given a homotopy from X to B $h : X \times [0,1] \to B$ and a lift $\tilde{h}_0 : X \to E$ lifting the starting point of the homotopy $h(0) \equiv h_0$ (i.e., $h_0 = p \circ \tilde{h}_0$), there exists a homotopy $\tilde{h} : X \times [0,1] \to E$ lifting h with $\tilde{h}(0) = \tilde{h}_0$. Diagrammatically:

$$\begin{array}{ccc} X \times \{0\} & \xrightarrow{\tilde{h}_0} & E \\ \downarrow & \overset{\tilde{h}}{\nearrow} & \downarrow p \\ X \times [0,1] & \xrightarrow{h} & B \end{array}$$

where the dashed arrow represents the induced map (i.e., the map we get by specifying the other maps). The homotopy lifting property thus tells us that specifying the lift of the starting point of a homotopy, we have a lift of the entire homotopy.

Fibrations A fibration is a map p (as above) that obeys the homotopy lifting property for all spaces X. One calls B the base space, E the total space, and $p^{(-1)}(b) \subset E$ the fiber over $b \in B$. From this nomenclature one sees that the fibration is a generalization of a fiber bundle, in the sense that fiber bundles obey the homotopy lifting property but fibrations don't rely on local trivializations and instead focus on homotopy. Occasionally fibrations are written $F \to E \to B$ for a fiber F.

3.3.1 Path space fibration

One important example of a fibration is the **path space fibration**. It is constructed as follows: consider the space of paths on a space B starting at some fixed point $b \in B$, denoted

$$PB = \{f : [0,1] \to B \mid f(0) = b\}$$

Futher, let $\pi : PB \to b$ be the projection onto the endpoint of the path: $\pi(f) = f(1)$.

Check it's a fibration This turns out to be a fibration in the following sense. Take some space X and a homotopy into B: $h_t : X \to B$ (I am using the Hatcher notation where one tucks the $[0,1]$ portion of the homotopy into a subscript). Say we're also given a lift $\tilde{h}_0 : X \to PB$, which we choose to write as $\tilde{h}_0(x) = (b, \gamma_x) \in PB$ for $f : X \to b \in B$, and $\gamma_x : I \to B$ (here this 2-tuple notation denotes a path γ_x starting at b).

The lift we construct is $\tilde{h}_t = (B, \gamma_x \circ h_{[0,t]}(x))$, where $\circ$ here denotes path composition. Thus, this is a path starting at b, following γ_x (the lift of h the homotopy from X to B), then following $h_s(x)$ along $s \in [0,t]$, ending at $h_t(x)$. Notice t here is the parameterization coming from the homotopy $h_t : X \to B$, and the paramterization of $\gamma_x \circ h_{[0,t]}(x)$ is that of PB:

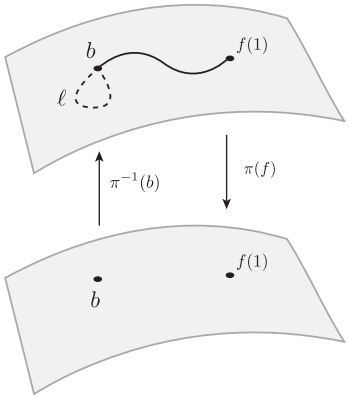
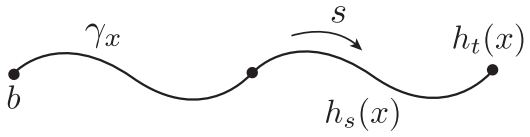


Figure 3.1: Sketch of the path space fibration.



Thus, this is a path in PB; then what is the projection $\pi(\tilde{h}_t)$, i.e., its endpoint? Well, that's simply $h_t(x)$. Thus $\tilde{h}_t$ is indeed a lift and we can see the sense in which it is a fibration.

What is the fiber $\pi^{-1}(b)$? It is the space of loops based at b. This is because $\pi^{-1}(b)$ are precisely the paths which must end at b and start at b. Thus we have a fibration $\Omega B \to PB \to B$. See Hatcher, Prop. 4.64 for a generalization of this proof.

3.4 Third Homotopy of Compact Lie Group is Z

3/9/24

It is the case that for any simple, compact Lie group G one has $\pi_3(G) = \mathbb{Z}$. Here I provide a sketch of a proof that this is the case.

Apply the path fibration construction to G:

$$\begin{array}{ccc} X \times \{0\} & \xrightarrow{\tilde{h}_0} & PG \\ \downarrow & \overset{\tilde{h}}{\nearrow} & \downarrow \pi \\ X \times [0,1] & \xrightarrow{h} & G \end{array}$$

with fiber ΩG the loopspace. Every fibration grants us a long exact sequence of homotopy groups; in this case we have

$$\cdots \to \pi_n(\Omega G) \to \pi_n(PG) \to \pi_n(G) \to \pi_{n-1}(\Omega G) \to \pi_{n-1}(PG) \to \cdots$$

But PG is contractible (they're just paths; while you might be worried that loops are not contractible, this is not the sense in which we are considering contraction. Here the appropriate way to think about it is as a line being "sucked" back into its starting point), so all homotopies vanish and in fact one has

$$\cdots \to \pi_n(\Omega G) \to 0 \to \pi_n(G) \to \pi_{n-1}(\Omega G) \to 0 \to \cdots \tag{3.1}$$

Which implies that $\pi_n(G) = \pi_{n-1}(\Omega G)$, and it suffices to compute $\pi_{n-1}(\Omega G)$. This is an easier task because it is possible to approximate loops by geodesics, and geodesics of Lie groups are relatively nice to work with. In particular, we may approximate a loop based at the origin e by k geodesics, all of which we take to be emanating from the origin of a copy of G:

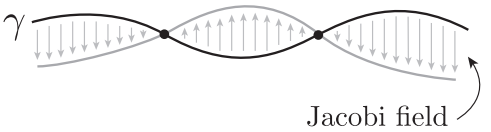
$$\Omega G \approx \text{geodesics of } S = \text{open subset of } G \times \overset{k}{\cdots} \times G$$

where one can take better and better approximations by cranking up k. We are able to think of such a loop of broken geodesics as living in k copies of G because G's metric is biinvariant, a fact which is derived from its compactness.

Now one considers the energy functional $E(\gamma) = (\int |\gamma|^2)$, for which the geodesic is, by definition, a critical point (ideally, it minimizes it, but this is a subtle matter). Conjugate points on this geodesics are points such that any Jacobi field along γ keeps those points fixed (i.e., it vanishes at that point); alternatively, they are points such that there exists a family of geodesics keeping those points fixed. The Morse index theorem then tells us that the number of such conjugate points is equal to the "index" of the Hessian of $E(\gamma)$.

Breaking this statement down: the Hessian is essentially a matrix of second derivatives; it characterizes the critical point at which we are located. Namely, the number of negative and positive eigenvalues characterizes the type of critical point; e.g., for n_- negative eigenvalues and n_+ positive eigenvalues one is at a saddle point with $2n_-$ "decreasing directions". n_- is the precisely the index of the Hessian.

Thus we gain some geometric intuition for why the number of conjugate points of the geodesic is equal to the index: let's say we have a geodesic γ related by a Jacobi field to another geodesic γ' such that there are n conjugate points. The fact that γ' is also a geodesic means that we can't possibly be at a minimum of the energy functional, though we are at a critical point. This is only possible if there the energy functional decreases in certain directions in path space; there are as many such directions as there are sections of the geodesic pinned by conjuate points, as the conjugate points are necessarily shared as boundaries by a family of neighboring geodesics.



An additional two facts we need are (1) the index of geodesics of Lie groups is always even, and (2) the fundamental theorem of Morse theory: the space of paths from p to q is homotopic to a CW complex with one cell of dimension λ for each geodesic from p to q of index λ. For (1), one can compute this explicitly, but I don't find the computation particularly insightful, so I am putting the matter aside until I can think of a more intuitive way to see it (see here). For (2), I hope it is relatively unsurprising upon some thought—it is a special case of the fact that any manifold is a CW complex with a cell of dimension λ for each critical point of index λ (TODO understand this?).

So every geodesic has an even index, and the space of geodesics has an λ-cell for each geodesic of index λ. Because we can approximate loops by geodesics, the space of loops only has even cells, and thus $\pi_1(\Omega G) = 0$, and also, as a bonus, the cell complex has

$$\cdots \longrightarrow 0 \xrightarrow{\partial} \mathbb{Z}\{\text{2-cells}\} \xrightarrow{\partial} 0$$

(since there are no 1- or 3-cells), meaning that

$$H_2(S) = \frac{\ker \partial = \mathbb{Z}\{2-\text{cells}\}}{\operatorname{im} \partial = \{0\}} = \mathbb{Z}^t$$

for some t (the group is generated by t 2-cells). The Hurewicz theorem tells us for an n-connected space the next higher $(n+1)$ homotopy and homology groups agree; thus if G is path-connected, we have $\pi_2(S) = H_2(S) = \mathbb{Z}^t$. But we've argued that S can be made to be a good enough approximation of ΩG; thus, we have $\pi_2(\Omega G) = \mathbb{Z}^t$ and $\pi_3(G) = \mathbb{Z}^t$ since $\pi_3(G) \simeq \pi_2(\Omega G)$ by eq. (3.1).

Now t here is related to the number of simple Lie groups that make G up. Thus, if G is simple we have $t = 1$, which is what we want.

References

- Matt Noonan, Homotopy groups of Lie groups, URL (version: 2009-12-15): https://mathoverflow.net/q/8957
- Milnor, Morse Theory (1960).

3.5 2+1 Electrodynamics ...

3.6 Dirac Quantization and Periodic Lagrange Multipliers 3/15/24

Consider a Dirac-quantization-like constraint on a field ϕ.

$$\int_{\partial S} f[\phi] \in \mathbb{Z} \tag{3.2}$$

Then consider some Lagrange multipler field $\sigma(x)$ enforcing $df[\phi] = 0$. The path integral reads

$$\int \mathcal{D}\phi\mathcal{D}\sigma \exp\left(i \int \mathcal{L}[\phi] + \sigma df[\phi]\right) \tag{3.3}$$

Consider what happens if we shift $\sigma \to \sigma + 2\pi$:

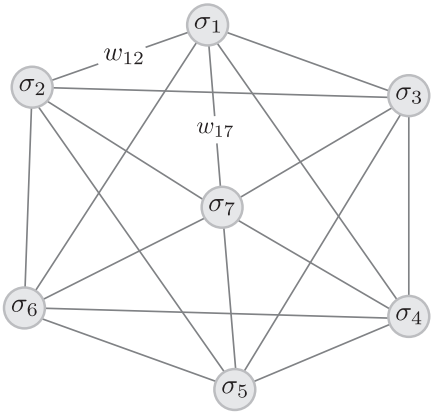
$$\int \mathcal{D}\phi\mathcal{D}\sigma \exp\left(i \int \mathcal{L}[\phi] + \sigma df[\phi] + 2\pi df[\phi]\right)$$

Now apply Stokes' theorem and eq. (3.2) to the last term:

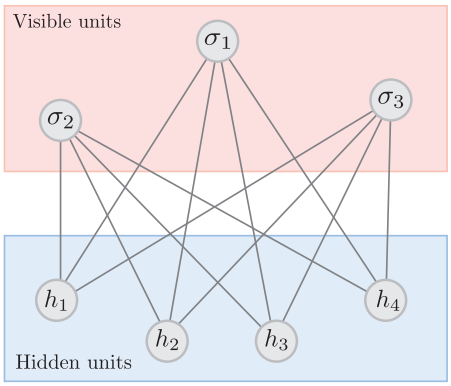
$$\begin{aligned}&\int \mathcal{D}\phi\mathcal{D}\sigma \exp\left(i \int \mathcal{L}[\phi] + \sigma df[\phi] + i\int_{\partial S} 2\pi f[\phi]\right)\\ &\quad= \int \mathcal{D}\phi\mathcal{D}\sigma \exp\left(i \int \mathcal{L}[\phi] + \sigma df[\phi] + 2\pi i N\right), \qquad N \in \mathbb{Z}\end{aligned}$$

But of course we can just extract $\exp(2\pi i N)$ and set it to unity, so the path integral returns to the form eq. (3.3). Thus, as seen from the perspective of the path integral $\sigma \sim \sigma + 2\pi$.

any space z is whe to cell complex. whe $\implies$ h.e. by whitehead



(a) Boltzmann machine.



(b) Restricted Boltzmann machine. h_i do not connect to other h_i and likewise for σ_i.

Figure 3.2

3.7 Instantons and Confinement ...

3.8 Continuum Limit of Gutzwiller Wavefunctions

3.9 Bosonization in 2+1 ...

3.10 Neural Quantum States 3/16/24

3.10.1 (Restricted) Boltzmann Machines

Boltzmann machines are a type of neural network which leverage insights from statistical physics to perform unsupervised learning of probabilistic models. That is, the goal is to use the Boltzmann machine to learn some target probability distribution $P_{\mathrm{T}}(X)$ given samples drawn from that probability distribution.

The machine itself consists of a set of units σ_i and weights w_{ij} connecting each of those units (fig. 3.2a). One defines an energy function

$$E(\sigma; w_{ij}, a_i) = \sum_i a_i \sigma_i + \sum_{ij} \sigma_i w_{ij} \sigma_j \tag{3.4}$$

with biases a_i and weights w_{ij} connecting each unit to every other unit. The units themselves are binary with ± 1 (some other authors choose $\{0, 1\}$). This model belongs to the class of **energy-based models**; such models leverage the fact that for an appropriately chosen $E(v)$ (in the Boltzmann machine case, appropriately chosen w_{ij}) the target probability distribution—call this $p_{\mathrm{T}}(\sigma)$—can be sufficiently approximated by the Boltzmann distribution

$$p(\sigma) = \frac{1}{Z} e^{-E(\sigma; w_{ij}, a_i)}$$

where

$$Z = \sum_{\{\sigma_i\}} \exp(-E(\sigma_i; w_{ij}, a_i))$$

is the partition function, i.e., the sum over all possible σ_i configurations that functions as a normalizing constant.

The model then trains w_{ij} by minizing some sort of measure of the divergence of $p(\sigma)$ from $p_{\mathrm{T}}(\sigma)$—for instance maximum likelihood estimation or the Kullbach-Leibler divergence. Thus, given an unlabelled dataset $\{\boldsymbol{\sigma}_n\}$ which is drawn from the unknown distribution p_{T}, we may find $E(\sigma; w_{ij}, a_i)$ such that $p(\sigma)$ approximates p_{T}.

In practice, however, it is quite difficult to sample $p(\sigma)$ for training purposes. The culprit is Z, whose difficulty to evaluate is well-known to physicists. One then resorts to the usual techniques employed to evaluate complex sums or integrals of this kind—chiefly, Monte Carlo methods. We will see that the training of neural quantum states is no different and we will need to employ Monte Carlo sampling of the wavefunction.

A special case of the Boltzmann machine that was designed with training convenience in mind is the **restricted Boltzmann machine**. A restricted Boltzmann machine demotes a subset of σ_i to hidden units denoted h_i—these are units which do not appear as arguments to the probability distribution $p(\sigma_i)$ and whose purpose is analogous to a hidden layer in a feedforward network. Further, and crucially, we stipulate that the hidden units do not connect to themselves, only to visible units, and likewise for visible units (fig. 3.2b). This amounts to writing $\sigma_i w_{ij} \sigma_j$ as $h_i w_{ij} \sigma_j$, so that

$$E(\sigma_i; w_{ij}, a_i, b_i) = \sum_i a_i \sigma_i + \sum_i b_i h_i + \sum_{ij} h_i w_{ij} \sigma_j$$

(we have separated the bias into a visible bias a_i and a hidden bias b_i). The target probability distribution is approximated by the marginalization over h_i—the trace, in physics parlance—so we must sum over h_i

$$p(\sigma_i) = \sum_{\{h_i\}} e^{-E(\sigma, h)}$$

The gradient (of the log) for the n-th vector $\boldsymbol{\sigma}_n$

$$\frac{\partial \log p(\hat{\boldsymbol{\sigma}}_n)}{\partial w_{ij}} = \frac{1}{p(\hat{\boldsymbol{\sigma}}_n)} \frac{\partial p(\hat{\boldsymbol{\sigma}}_n)}{\partial w_{ij}} = \frac{1}{p(\hat{\boldsymbol{\sigma}}_n)} \left[\sum_{\{h_i\}} e^{-E(v,\hat{\boldsymbol{\sigma}}_n)} \frac{\partial}{\partial w_{ij}} \left(\frac{1}{Z} \right) + \frac{1}{Z} \sum_{\{h_i\}} \frac{\partial}{\partial w_{ij}} e^{-E(h,\hat{\boldsymbol{\sigma}}_n)} \right]$$

$$= \frac{1}{Z} \frac{1}{p(\hat{\boldsymbol{\sigma}}_n)} \left[- \sum_{\{h_i\}} e^{-E(v,\hat{\boldsymbol{\sigma}}_n)} \frac{1}{Z^2} \frac{\partial Z}{\partial w_{ij}} - \frac{1}{Z} \sum_{\{h_i\}} e^{-E(h,\hat{\boldsymbol{\sigma}}_n)} \hat{\sigma}_{n,i} h_j \right]$$

The derivative of the partition function is

$$\frac{\partial Z}{\partial w_{ij}} = \frac{\partial}{\partial w_{ij}} \sum_{\{h_i\}} \sum_{\{\sigma_i\}} e^{-E(h,\boldsymbol{\sigma}_n)} = \sum_{\{h_i\}} \sum_{\{\sigma_i\}} e^{-E(h,\boldsymbol{\sigma}_n)} (-\sigma_i, h_j) = -Z \langle \sigma_i v_j \rangle_{\text{model}}$$

The subscript "model" says that this is the expectation value of the current forward pass Boltzmann machine distribution as opposed to the expectation value with respect to the data, which will show up shortly. Then we have

$$\frac{\partial \log p(\hat{\boldsymbol{\sigma}}_n)}{\partial w_{ij}} = \frac{1}{p(\hat{\boldsymbol{\sigma}}_n)} \left[\overbrace{\frac{1}{Z} \sum_{\{h_i\}} e^{-E(v,\hat{\boldsymbol{\sigma}}_n)}}^{p(\hat{\boldsymbol{\sigma}})} \frac{Z}{Z} \langle \sigma_i h_j \rangle_{\text{model}} - \overbrace{\frac{1}{Z} \sum_{\{h_i\}} e^{-E(h,\hat{\boldsymbol{\sigma}}_n)}}^{p(\hat{\boldsymbol{\sigma}})} \hat{\sigma}_{n,i} h_j \right]$$

$$= \langle \sigma_i h_j \rangle_{\text{model}} - \hat{\sigma}_{n,i} h_j$$

Now, the gradient is the expected value over the training vectors $\hat{\boldsymbol{\sigma}}_n$, so the derivative is

$$\frac{\partial \log p(\hat{\boldsymbol{\sigma}}_n)}{\partial w_{ij}} = \langle \sigma_i h_j \rangle_{\text{model}} - \langle \hat{\sigma}_{n,i} h_j \rangle_{\text{data}}.$$

The benefit of the restricted Boltzmann machine is that there is an exact distribution that h_j follows given some $\hat{\sigma}_n$, so we can evaluate the second term. Let's say that $h_k \in \{0, 1\}$; the probability that $h_k = 1$ given a visible unit can be calculated as follows. We have

$$\frac{p(h_k = 1 \mid \sigma)}{p(h_k = 0 \mid \sigma)} = \frac{\exp\left[b_k + \sum_{i \neq k h_i b_i} + \sum_i v_i a_i + \sum_{i \neq k,j} h_i w_{ij} v_j + \sum_j w_{kj} v_j \right]}{\exp\left[0 \cdot b_k + \sum_{i \neq k h_i b_i} + \sum_i v_i a_i + \sum_{i \neq k,j} h_i w_{ij} v_j + 0 \cdot \sum_j w_{kj} v_j \right]}$$

Now, since h_k can only be in two states, $p(h_k = 0 \mid \sigma) = 1 - p(h_k 0 \mid \sigma)$, and

$$\frac{p(h_k = 1 \mid \sigma)}{1 - p(h_k = 1 \mid \sigma)} = \frac{\exp\left[b_k + \sum_{i \neq k h_i b_i} + \sum_i v_i a_i + \sum_{i \neq k,j} h_i w_{ij} v_j + \sum_j w_{kj} v_j \right]}{\exp\left[\sum_{i \neq k h_i b_i} + \sum_i v_i a_i + \sum_{i \neq k,j} h_i w_{ij} v_j \right]}$$

$$= \exp\left(b_k + w_{kj} v_j \right)$$

We can work out

$$p = e^W(1-p) \implies (1+e^W) = e \implies p = \frac{e^W}{1+e^W} \implies \frac{1}{1+e^{-W}} = \sigma(-W)$$

So that the above implies

$$p(h_k = 1 \mid \sigma) = \sigma(b_k + \sum_j w_{kj} v_j).$$

Notice that the argument relies on the fact that h_i does not connect to other hidden units. An identical argument finds

$$p(\sigma_k = 1 \mid h) = \sigma(a_k + \sum_j w_{jk})$$

Sampling $\langle \sigma_i h_j \rangle_{\text{model}}$, however, remains difficult. Strategies for doing so that leverage $p(h_k = 1 \mid \sigma)$ and $p(\sigma_k = 1 \mid h)$ are described in Hinton. In the case of neural quantum states, however, the natural loss function is the energy of the system, and the gradients we must calculate are not the above. The evaluation will still be difficult, but in this case only because we must obtain the energy and its gradients indirectly in order to circumvent the curse of dimensionality.

3.10.2 Neural Quantum States

One can view the many-body wavefunction $\psi(\sigma_1, \ldots, \sigma_N)$ appearing in

$$|\psi\rangle = \sum_{\{\sigma_i\}} \psi(\sigma_1, \ldots, \sigma_N) \, |\sigma_1, \cdots, \sigma_N\rangle$$

as a function from local degrees of freedom to a complex number, the amplitude, $\psi : \{\sigma_i\} \to \mathbb{C}$. Thus we may treat ψ as a a Boltzmann machine with visible units $\{\sigma_i\}$, albeit one that produces a complex number corresponding to the wavefunction as opposed to the probability amplitude. We thus provide ψ with a set of weights $W = (w_{ij}, a_i, b_i)$ and hidden units h_i such that

$$\psi(\sigma_1, \ldots, \sigma_N; W) = \sum_{\{h_i\}} \exp\left(\sum_j a_j \sigma_j^z + \sum_i b_i h_i + \sum_{ij} w_{ij} h_i \sigma_j^z \right)$$

Because ψ is complex-valued, w_{ij}, a_i, b_i are likewise taken to be complex valued. This is a restricted Boltzmann machine; we thus can explicitly marginalize ψ:

$$\begin{aligned}\psi &= \sum_{\{h_i\}} \exp\left(\sum_j a_j \sigma_j^z + \sum_i b_i h_i + \sum_{ij} w_{ij} h_i \sigma_j^z\right) \\ &= \sum_{\{h_i\}} \exp\left(\sum_j a_j \sigma_j^z\right) \prod_i^M \exp\left(\sum_i b_i h_i\right) \exp\left(\sum_j w_{ij} h_i \sigma_j^z\right)\end{aligned}$$

where we move the summation into the product as follows:

$$\begin{aligned}&= \exp\left(\sum_j a_j \sigma_j^z\right) \prod_i^M \sum_{h_i=\{\pm 1\}} \exp\left(\sum_i b_i h_i\right) \exp\left(\sum_{ij} w_{ij} h_i \sigma_j^z\right) \\ &= \exp\left(\sum_j a_j \sigma_j^z\right) \prod_i^M \left[\exp\left(b_i + \sum_j w_{ij} \sigma_j^z\right) + \exp\left(-b_i - \sum_j w_{ij} \sigma_j^z\right)\right] \\ &= \exp\left(\sum_j a_j \sigma_j^z\right) \prod_i^M 2\cosh\left(b_i + \sum_j w_{ij} \sigma_j^z\right)\end{aligned}$$

Note

For some reason I always have to convince myself that the above move works, so for future reference here's the argument: we have a product indexed by i and a sum indexed by ℓ, which expanded looks like

$$\begin{aligned}\prod_i^I \sum_\ell^L a_{i\ell} = (a_{11} + a_{12} + \cdots + a_{1L}) \\ \times (a_{21} + \cdots + a_{2L}) \\ \cdots \times (a_{I1} + \cdots + a_{IL})\end{aligned}$$

If we expand the products of sums, we obtain

$$= a_{11}a_{21}\cdots a_{I1} + \cdots + a_{1L}a_{2L}\cdots a_{IL} = \sum_{a_{1i}}^L \sum_{a_{2i}}^L \cdots \sum_{a_{Li}}^L \prod_i^I a_{i\ell}.$$

Thus

$$\psi(\sigma_1,\ldots,\sigma_N;W)=\exp\left(\sum_j a_j\sigma_j^z\right)\prod_i^M 2\cosh\left(b_i+\sum_j w_{ij}\sigma_j^z\right)$$

Given a Hamiltonian H, how do we update w_{ij}, a_i, b_i to minimize the energy $H=\langle\psi|\,H\,|\psi\rangle$? As is typical with this type of problem, one resorts to a Monte Carlo method. Carleo and Troyer use a technique called stochastic reconfiguration wherein one updates the weights as

$$W_{k+1,i}=W_{k,i}-\gamma_k S^{-1}_{k,ij}F_{k,j}\qquad (k=\text{iteration step})$$

where

$$S_{k,ij}=\langle O_i^\dagger O_j\rangle-\langle O_i^\dagger\rangle\langle O_j\rangle,\qquad F_i=\langle E_{\text{loc}}O_i^\dagger\rangle-\langle E_{\text{loc}}\rangle\langle O_i^\dagger\rangle$$
$$O_i=\frac{1}{\psi}\partial_{W_i}\psi,\qquad E_{\text{loc}}=\frac{1}{\psi}\langle\{\sigma_i\}|\,H\,|\psi\rangle$$

Here the expectation values denote $\langle A\rangle=\sum_{\{\sigma_i\}}A(\{\sigma_i\})|\psi(\{\sigma_i\})|^2$. They are evaluted with a Markov chain sampled with a Metropolis-Hastings procedure: as is common with these VMC techniques, we flip a random spin and accept the new configuration with probability

$$P(\{\sigma_i\}_k\to\{\sigma_i\}_{k+1})=\min\left(1,\left|\frac{\psi(\{\sigma_i\}_{k+1})}{\psi(\{\sigma_i\}_k)}\right|^2\right).$$

Notice that we do not need to evaluate any Slater determinants and the evaluation of the Metropolis-Hastings weight is fairly straightforward. One caveat is that S is not guaranteed to be nonsingular; Carleo and Troyer circumvent this with an explicit regularization

$$S^{\text{reg}}_{ij}=S_{ij}+\lambda_k\delta_{ij}S_{ij},\qquad \lambda_k=\max(\lambda_0 b^k,\lambda_{\min})$$

where their parameters are $\lambda_{\min}=10^{-4}$, $\lambda_0=100$, $b=0.9$.

Stochastic reconfiguration is known in the machine learning literature as natural gradient descent; there, what we call S is playing the role of the Fisher information matrix, and F plays the role of the gradient. In that sense it is a so-called "preconditioning" on the gradient as it takes the form of an adjustment to the gradient (the multiplication byt he inverse of the Fisher matrix).

References

- Giuseppe Carleo, Matthias Troyer, *Solving the quantum many-body problem with artificial neural networks.* Science 355,602-606 (2017). DOI:10.1126/science.aag2302. See the supplemental materials for information on their Monte Carlo sampling.

- Torlai, Giacomo, and Roger G. Melko. *"Learning thermodynamics with Boltzmann machines."* Physical Review B 94.16 (2016): 165134. DOI:10.1103/PhysRevB.94.165134.
- Grathwolh, et al., *Your Classifier is Secretly an Energy Based Model and You Should Treat it Like One* (2019). arXiv:1912.03263.
- Hinton, *A Practical Guide to Training Restricted Boltzmann Machines*. https://www.cs.toronto.edu/~hinton/absps/guideTR.pdf.
- Sorella, *Green Function Monte Carlo with Stochastic Reconfiguration*, DOI:10.1103/PhysRevLetter.80.5668.

3.11 Poincaré Duality, Particle-Vortex, & All That 3/17/24

3.11.1 1 + 1

Bosonization

$$\int \mathcal{D}\psi \exp i\left(\int d^2x\left[-\bar{\psi}\not{\partial}\psi + i\bar{\psi}\not{A}\psi\right]\right) = \exp\left(\frac{i}{4\pi}\int d^2x F^{\mu\nu}\frac{1}{\partial^2}F_{\mu\nu}\right) \tag{3.5}$$

The left-hand side is of course Gaussian, so we get the usual fermionic determinant:

$$\int \mathcal{D}\psi \exp\left(\int d^2x - \bar{\psi}(i\not{\partial} + \not{A})\psi\right) = \det(i\not{\gamma} + \not{A})$$

How do we evaluate this determinant? There is an exercise in Jackiw, "Topological Investigations of Quantized Gauge Theories" which provides a 1d sketch of how to do so. The story is like this: One thinks of this determinant as the effective energy

$$W[A] = -i\ln\det(\not{\partial} - \not{A}), \qquad e^{iW[A]} = \langle 0|0\rangle_A$$

(that is, we are viewing the path integral as the vacuum-vacuum correlation function) so that we may substitute the Green's function for $\not{\partial}$:

$$(\not{\partial} - i\not{A})G = \delta(x) - iG\not{A}$$

Using $\det M = \exp\ln M$ we write

$$\ln\det(\not{\partial} - i\not{A}) = \ln\exp(\operatorname{tr}\ln(\delta - iG\not{A})) = \operatorname{tr}\ln(\delta - iG\not{A})$$

One expands the log and finds that only the first order term contributes. (I wasn't sure how this was the case. Perhaps it had something to do with the step functions.) Then we substitue G for $i\langle\text{propagator}\rangle$ and find

$$-i\operatorname{tr} Ge\not{A} = \int dtA(t)$$

In $1+1$d one supposedly finds

$$\int dxdy A^\mu(x) G_{\mu\nu} A^\nu(y)$$

with $G_{\mu\nu} = [g_{\mu\nu} - \partial_\mu\partial_\nu/\partial^2]\delta(x-y)$ for the vector current propagator. This is a classic result owing to the famed Schwinger paper in which he introduces what we now call the Schwinger model. At any rate I could not really piece together the story for $1+1$, and even taking for granted the above equation I could not get the form in eq. (3.5). One thought was to leverage $\partial^2 = g^{\mu\nu}\partial_\mu\partial_\nu \implies g_{\mu\nu} = \partial_\mu\partial_\nu/\partial^2$ but I was still far from the desired result.

T-duality

3.11.2 $2+1$

Particle-vortex Consider the abelian Higgs model:

$$S = \int d^3x \left[-\frac{1}{2}|\mathcal{D}_\mu\phi|^2 - V(|\phi|) - \frac{1}{4}F_{\mu\nu}^2 \right]$$

The field ϕ admits vortex configurations $\phi = \phi_0(r)\exp(i\alpha(\theta))$ where $\alpha(\theta) = N\theta$ and N is the winding number of the vortex. We split the field configuration into vortex solutions and smooth solutions $\phi = \phi_0 \exp(i\alpha_{\mathrm{V}} + i\alpha_{\mathrm{S}})$, so that the kinetic term becomes

$$\begin{aligned}
|\mathcal{D}_\mu\phi|^2 &= |\partial_\mu\phi_0 \cdot e^{i\alpha} + \phi_0\partial_\mu e^{i\alpha_{\mathrm{V}}+i\alpha_{\mathrm{S}}} + iea_\mu\phi_0 e^{i\alpha}|^2 \\
&= |\partial_\mu\phi_0 \cdot e^{i\alpha} + i\phi_0 e^{i\alpha}(\partial_\mu\alpha_{\mathrm{V}} + \partial_\mu\alpha_{\mathrm{S}}) + iea_\mu\phi_0 e^{i\alpha}|^2 \\
&= |\partial_\mu\phi_0 \cdot e^{i\alpha} + i\phi_0 e^{i\alpha}(\partial_\mu\alpha_{\mathrm{V}} + \partial_\mu\alpha_{\mathrm{S}} + ea_\mu)|^2 \\
&= \underbrace{|e^{i\alpha}|^2}_{=1} |\partial_\mu\phi_0 + i\phi_0(\partial_\mu\alpha_{\mathrm{V}} + \partial_\mu\alpha_{\mathrm{S}} + ea_\mu)|^2
\end{aligned}$$

All the quantities involved in the second factor are real so this is just the square modulus of a complex number:

$$\mathcal{D}_\mu\phi = |\partial_\mu\phi_0|^2 + \phi_0^2\left(\partial_\mu\alpha_{\mathrm{V}} + \partial_\mu\alpha_{\mathrm{S}}\right)^2$$

Note that because the curl of a gradient vanishes, we have $\epsilon^{ij}\partial_i\partial_j\alpha = 0$ everywhere except at the vortices. We use the shorthand $b_\mu = \partial_\mu\alpha$ and enforce the aforementioned constraint (now reading $\epsilon^{\mu\nu\rho}\partial_\mu b_\nu = 0$) via a Lagrange multiplier field λ_μ. The action now reads

$$S = \int d^3x \left[-\frac{1}{2}|\partial_\mu\phi_0|^2 - \frac{1}{2}\phi_0^2\left(b_{\mu,\mathrm{S}} + b_{\mu,\mathrm{V}} + ea_\mu\right)^2 + \lambda_\mu\epsilon^{\mu\nu\rho}\partial_\nu b_{\rho,\mathrm{S}} + \cdots \right] \tag{3.6}$$

The equation of motion of $b_{\mu,\mathrm{S}}$ is

$$\frac{\partial\mathcal{L}}{\partial b_{\mu,\mathrm{S}}} - \partial_\nu\frac{\partial\mathcal{L}}{\partial_\nu b_{\mu,\mathrm{S}}} = 0 \tag{3.7}$$

There's some index rearranging to peform to obtain the second term: the derivative appears in $\lambda_\mu \epsilon^{\mu\nu\rho}\partial_\nu b_{\rho,\mathrm{S}}$ which we can rewrite as $\lambda_\rho \epsilon^{\rho\nu\mu}\partial_\nu b_{\mu,\mathrm{S}} = -\lambda_\rho \epsilon^{\mu\nu\rho}\partial_\nu b_{\mu,\mathrm{S}}$ by relabelling indices and rearranging the antisymmetric tensor. We then have

$$\partial_\nu \frac{\partial \mathcal{L}}{\partial_\nu b_{\mu,\mathrm{S}}} = \partial_\nu \frac{\partial}{\partial_\nu b_{\mu,\mathrm{S}}} \left(\lambda_\rho \epsilon^{\mu\nu\rho}\partial_\nu b_{\mu,\mathrm{S}}\right) = \partial_\nu \lambda_\rho \epsilon^{\mu\nu\rho}$$

Meanwhile,

$$\phi_0^2 \left(b_{\mu,\mathrm{S}} + b_{\mu,\mathrm{V}} + e a_\mu\right)^2 = \phi_0^2 \left(b_{\mu,\mathrm{V}}^2 + b_{\mu,\mathrm{S}}^2 + e^2 a_\mu^2 + 2 b_{\mu,\mathrm{V}} b_{\mu,\mathrm{S}} + 2 b_{\mu,\mathrm{V}} e a_\mu + 2 b_{\mu,\mathrm{S}} e a_\mu\right)$$

so

$$\frac{\partial \mathcal{L}}{\partial b_{\mu,\mathrm{S}}} = -\phi_0^2 (b_{\mu,\mathrm{S}} + b_{\mu,\mathrm{V}} + e a_\mu)$$

and eq. (3.7) is

$$\phi_0^2 (b_\mu + e a_\mu) = \epsilon^{\mu\nu\rho}\partial_\nu \lambda_\rho \iff b_\mu = \frac{\epsilon^{\mu\nu\rho}\partial_\nu \lambda_\rho}{\phi_0^2} - e a_\mu$$

Now we substitute this into eq. (3.6) to eliminate $b_{\mu,\mathrm{S}}$ and find the dual action. First we find

$$\phi_0^2 \left(b_{\mu,\mathrm{S}} + b_{\mu,\mathrm{V}} + e a_\mu\right)^2 = \frac{1}{\phi_0^2} (\epsilon^{\mu\nu\rho}\partial_\nu \lambda_\rho)^2 = \frac{1}{2\phi_0^2} (\partial_\mu \lambda_\nu - \partial_\nu \lambda_\mu)^2.$$

Using

$$b_{\mu,\mathrm{S}} = \frac{\epsilon^{\mu\nu\rho}\partial_\nu \lambda_\rho}{\phi_0^2} - b_{\mu,\mathrm{V}} - e a_\mu$$

we can substitute into the Lagrange multiplier term to find

$$\epsilon^{\mu\nu\rho}\lambda_\mu \partial_\nu b_{\rho,\mathrm{S}} = \frac{1}{\phi_0^2} \lambda_\mu \epsilon^{\mu\nu\rho}\epsilon^{\rho\lambda\sigma}\partial_\nu \partial_\lambda \lambda_\sigma - \epsilon^{\mu\nu\rho}\lambda_\mu \partial_\nu b_{\sigma,\mathrm{V}} - e\epsilon^{\mu\nu\rho}\lambda_\mu \partial_\nu a_\mu$$

where supposedly the first term vanishes. We recognize the second term as the vortex current

$$j^\mu_{\mathrm{vortex}} = \frac{e}{2\pi}\epsilon^{\mu\nu\rho}\partial_\nu b_\rho = \frac{1}{2\pi\phi_0^2}\epsilon^{\mu\nu\rho}\partial_\nu j_\rho$$

where j_ρ is the U(1) Noether current. This current is conserved by virtue of the (topological) fact that the curl of a divergence should vanish; thus it is an example of a topological current, one whose conservation is ensured by topology of the field configuration. One might think it funny to view it as a current, seeing as as which symmetry it is associated with is unclear. It is arguably this very question—what symmetry this current is associated with—that motivates the introduction of generalized symmetries. For instance, if we consider U(1) theory, invariance under a 1-form symmetry characterizes a lack of vortices (which are

charged under this symmetry). In fact, it is often the case that the generalized symmetry in the original theory is understood as a conventional symmetry in the dual theory.

At any rate, the dual action is

$$S = \int d^3x \left[-\frac{1}{4\phi_0^2}(\partial_\mu \lambda_\nu - \partial_\nu \lambda_\mu)^2 - \frac{1}{2}\epsilon^{\mu\nu\rho}\lambda_\mu \partial_\nu b_{\rho,\mathrm{V}} - e\epsilon^{\mu\nu\rho}\lambda_\mu \partial_\nu a_\mu + V(\phi_0) + F_{\mu\nu}^2 \right].$$

Now note that in eq. (3.6) ϕ_0^2 plays the role of the kinetic coupling for $(\mathcal{D}_\mu \alpha)^2$ whereas in the dual action the b field has $1/\phi_0^2$; thus this is indeed a weak/strong duality.

3.11.3 Poincaré Duality

Current is some charge distirbution dirac delta, then hodge dual to other current

d - n topological current (some defect in an otherwise smooth function) is dual to original n funciton.

References

- Jackiw, "Topological Investigations of Quantized Gauge Theories," *Current Algebra and Anomaly.*
- Năstase, *String Theory Methods for Condensed Matter Physics*, Ch. 21. Note Năstase uses b_μ for the Lagrange multipliers and λ_μ for $\partial_\mu \alpha$. Lagrange multipliers are usually denoted λ so I switched his notation.

3.12 Fast Computation of VMC Determinants 3/18/24

In a Monte Carlo simulation of a strongly correlated system one needs the ratio

$$\frac{\langle x'|\Phi\rangle}{\langle x|\Phi\rangle}$$

where x' is a proposed classical configuration and x is the present configuration. Of course, given a set of single-particle wavefunctions, the many-body wavefunction $\langle x|\Phi\rangle$ is some Slater determinant of the form

$$\det U = \begin{bmatrix} \phi_1(x_1) & \phi_2(x_1) & \cdots & \phi_N(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & & \phi_N(x_2) \\ \vdots & & \ddots & \\ \phi_1(x_N) & \phi_2(x_N) & & \phi_N(x_N) \end{bmatrix}$$

(we ignore the spin d.o.f.; reintroducing it simply amounts to a product of Slater determinants for up and down spins). We have made the arbitrary choice to index the wavefunction

along the columns and the particle along the rows; of course the other option is obtained by a transpose. Now, calculating this determinant is at best an $O(N^3)$ task—assuming that the number of particles is an extensive quantity we will quickly run into unfeasibly long simulation times for more than a few decimal places' accuracy on a larger than 4×4 lattice.

The cleverer way to go about this is to tabulate from the get-go all possible weights that we might encounter. Fortunately, this is not a difficult thing to calculate; the first thing we need is a matrix of the same form as above, only with a row with every possible site:

$$\det \tilde{U} = \begin{bmatrix} \phi_1(r_1) & \phi_2(r_1) & \cdots & \phi_N(r_1) \\ \phi_1(r_2) & \phi_2(r_2) & & \phi_N(r_2) \\ \vdots & & \ddots & \\ \phi_1(r_L) & \phi_2(r_N) & & \phi_N(r_N) \end{bmatrix}$$

Observe that one can recover the slater matrix U by selecting those rows of $\tilde{U}$ that correspond to occupied sites. As it turns out, the following matrix will contain the weights

$$W = \tilde{U}U^{-1}, \qquad W_{K,\ell} = \frac{\det U'}{\det U} = \frac{\langle x'|\Phi\rangle}{\langle x|\Phi\rangle} \tag{3.8}$$

where x' differs from x by the particle x_ℓ hopping to the site r_K. Here U is $N \times N$ matrix, $\tilde{U}$ is $L \times N$ (L is the total number of sites), and thus W is $L \times N$. Of course, upon accepting x', W changes by virtue of the appearance of U, which depends on the configuration. Then the matrix W can be updated as

$$W'_{I,j} = W_{I,j} - \frac{W_{I,\ell}}{W_{K,\ell}}(W_{K,j} - \delta_{\ell,j}), \qquad (x_\ell \to r_K)$$

Here $\delta_{\ell,j}$ is essentially the zero matrix with the ℓ-th column filled with ones. If only one particle is moved, this can be written as a rank-1 (dyadic) update

$$W'_{I,j} = W_{I,j} + a_I b_j, \qquad a_I = W_{I,\ell}, \quad b_j = -\frac{1}{W_{K,\ell}}(W_{K,j} - \delta_{\ell,j})$$

For implementation it is easier to think of this as the outer product

$$W' = W + \mathbf{a}\mathbf{b}^T$$

where it is crucial that we take the transpose and not the complex conjugate. Of course round-off errors might accumulate and at long simulation times this updating is susceptible to numerical instability. It is thus customary to occasionally recalculate W directly from the Slater matrix with eq. (3.8); at 64-bit precision this recalculation takes place on the order of a few million timesteps. Roundoff errors might also lead to large fluctiations on account of the matrix inverse of eq. (3.8)—an error of order ϵ of course becomes $1/\epsilon$ and one observes disproportionally large weights which do not affect the acceptance probability (at that timestep) but contaminate other matrix entries. In such situations it is necessary to round off elements of W at, say, roughly 15 decimal points for 64-bit precision floats.

References

- Becca, Sorella, *Quantum Monte Carlo Methods for Strongly Correlated Systems* (2017).

3.13 Local Unitary Transformations and Phase Transitions 3/26/24

Does symmetry local unitary transformation still connect two states in same phase iff there's no gap closing?

References

- Chen, Gu, Wen, *Local unitary transformation, long-range quantum entanglement, wave function renormalization, and topological order*. https://arxiv.org/abs/1004.3835.

3.14 Stationary Phase and Time Evolution in QM 3/27/24

3.15 SPTs ...

3.15.1 MPS Refresher

Matrix product states (MPS) take the form

$$|\psi\rangle = \sum_{\{j_i\}} A^{(j_1)}{}_{\mu_1} A^{(j_1)\mu_1}{}_{\mu_2} A^{(j_1)\mu_3}{}_{\mu_3} \cdots A^{(j_{n-1})\mu_{n-1}}{}_{\mu_{n-1}} A^{(j_n)\mu_n} |j_1 \cdots j_n\rangle \tag{3.9}$$

where $A^{(j_i)\mu_i}{}_{\mu_i}$ are $\chi \times \chi$ matrices for bond size χ (viewed as rank-3 tensors if we include the additional index j_i). They are ansatze wavefunctinos which leverage the local nature of entanglement to model to great approximation ground states of gapped systems. DMRG is for instance nothing more than a variational technique that explores the space of MPSs.

How do MPS leverage local entanglement struture? Consider some general state

$$|\psi\rangle = \sum_{\{j_i\}} \psi(j_1, \ldots, j_n) \, |j_1, \ldots, j_n\rangle$$

here $|j_i\rangle$ is a basis labelling the site degree of freedom; ultimately we want to express the wavefunction in this basis. Now, we are guaranteed a Schmidt decomposition of any state:

$$|\psi\rangle = \sum_{\alpha} \Lambda^{(\alpha)} \, |\alpha_{\mathrm{L}}\rangle \, |\alpha_{\mathrm{R}}\rangle$$

The strategy is to proceed along the chain and successively perform Schmidt decompositions at each site. We will see that this affords us a physically justifiable (and thus, one hopes, sufficiently accurate) reduction in the number of parameters we need to express the wavefunction.

To illustrate, the first two sites will proceed as

$$|\psi\rangle = \sum_{\alpha^1} A^{(\alpha^1)} |\alpha_1^1\rangle |\alpha_{2\cdots n}^1\rangle = \sum_{\alpha^1,\alpha^2} A^{(\alpha^1)} A^{(\alpha^2)} |\alpha_1^1\rangle |\alpha_2^1\rangle |\alpha_{3\ldots n}^2\rangle$$

here $A^{(\alpha^i)}$ are the Schmidt coefficients for that particular site, and $\alpha^i_{1\ldots m-1}, \alpha^i_{m\ldots n}$ label the states in our Schmidt basis for the decomposition partitoning the sites $m, \ldots, n$. We can treat $A^{(\alpha_i)} = A^{(\alpha_i)\mu_i}_{\quad\mu_i+1}$ diagonal matrix without loss of generality, for reasons that will become clear shortly. The dimension of this matrix is the size of the Schmidt basis, i.e., there is a diagonal entry for each α_i. The number of labels α^1 will depend on the particular site in the chain; as we move along, the number of labels will grow (as the number of states needed to form a basis exponentially with the number of sites). The key approximation that MPSs make is to keep only the χ lowest Schmidt values—the idea is that entanglement should remain relatively localized ...

Thus one obtains a decomposition

$$\sum_{\{\alpha_i\}} \Lambda^{(\alpha_1)}{}_{\mu_1} \Lambda^{(\alpha_2)\mu_1}_{\quad\mu_2} \cdots \Lambda^{(\alpha_{n-1})\mu_{n-2}}_{\quad\mu_{n-1}} \Lambda^{(\alpha_n)\mu_{n-1}} |\alpha_1 \ldots \alpha_n\rangle$$

At this point we only need to express this in site basis j_i by means of a change of basis matrix $\Gamma^{(j_i)\nu_i}_{\quad\mu_i} = \langle j_{i,\mu} | \alpha_{i,\nu} \rangle$:

$$\sum_{\{j_i\}} \Lambda^{(j_1)}_{\nu_1} \Gamma^{(j_1)\nu_1}_{\quad\mu_2} \Lambda^{(j_2)\mu_2}_{\quad\nu_2} \Gamma^{(j_2)\nu_2}_{\quad\mu_3} \cdots \Lambda^{(j_n)\mu_n}_{\quad\nu_n} \Gamma^{(j_n)\nu_n} |j_1 \ldots j_n\rangle$$

If we call $\Lambda^{(j_n)\mu_i}_{\quad\nu_i} \Gamma^{(j_i)\nu_i}_{\quad\mu_{i+1}} = A^{(j_i)\mu_i}_{\quad\mu_{i+1}}$ we recover the form of eq. (3.9).

3.16 Useful Field Theory Formuale 3/27/24

$$T^{\mu\nu} = \frac{\delta \mathcal{L}}{\delta \partial_\mu \phi} \partial^\nu \phi - g^{\mu\nu} \mathcal{L}$$

$$\mathcal{H} = \frac{\delta \mathcal{L}}{\delta \partial_0 \phi} \partial^0 \phi - g^{00} \mathcal{L}$$

3.17 Avoided Crossings and Symmetry 3/28/24

There is lore in the many-body literature that eigenvalues of a system undergoing an adiabatic evolution only cross if the eigenstates have different symmetries. Here's more or less how this works. Consider a diagonalized Hamiltonian H perturbed by a term H'

$$H + H' = \begin{bmatrix} E_1 & 0 \\ 0 & E_2 \end{bmatrix} + \begin{bmatrix} 0 & \Delta \\ \Delta^* & 0 \end{bmatrix} = \begin{bmatrix} E_1 & \Delta \\ \Delta^* & E_2 \end{bmatrix}$$

(without loss of generality we take the perturbation to be completely off-diagonal). The spectrum of this Hamiltonian consists of

$$|\psi_\pm\rangle = \phi_\pm |E_1\rangle - |E_2\rangle, \qquad E_\pm = E_1 + E_2 \pm \sqrt{(E_1 - E_2)^2 + 4|\Delta|^2}$$
$$\phi_\pm = E_1 - E_2 \pm \sqrt{(E_1 - E_2)^2 + 4|\Delta|^2}$$

Observe $E_\pm$: since $|\Delta|^2$ is nonnegative, we can only have $E_+ = E_-$ if $E_1 = E_2$ and $\Delta = 0$. This is an avoided crossing; we think of Δ as depending on some adiabatically varied parameter such that if any two eigenvalues are well represented by an effective 2×2 Hamiltonian as above, the energy levels cannot cross for any value of Δ.

This argument relies, however, on the mixing between $|E_1\rangle$ and $|E_2\rangle$. Superselection rules prevent, however, (adiabatic) evolution from a state in one irrep to another; more precisely we cannot adiabatically evolve into a superposition of states with two different symmetries. Thus, in our effective two-level system, $|E_1\rangle$ and $|E_2\rangle$ must enjoy the same symmetry; if they do not, they cannot mix and the energy levels can cross. (Is this in contradiction with the relationship beween degeneracy and symmetry? Not if you can have degeneracies that don't descent from symmetries.)

Chapter 4

April

Contents

4.1 Is RG a semigroup? 4/1/24

4.2

Integer quantum hall effect - we have a harmonic oscillator portion of the solution (in direction of applied magnetic field?), the quantization of which is the Landau level ν. Notice that if the landau level is completely filled, we have a gap. The edge modes carry the current; there are ν different "channels" for the modes. One can use Landauer electron transport theory to find that for each such channel we have $e^2 V_H/h$ current.

4.3 EFT of the Hubbard Model 4/6/24

Consider the Hubbard Hamiltonian (a indexes spin)

$$H = -t\sum_{j=1}^{L}(c_{j,a}^{\dagger}c_{j+1,a} + c_{j+1,a}^{\dagger}c_{j,a}) + \sum_{j=1}^{L} u(1-2n_{j\uparrow})(1-2n_{j\downarrow})$$

One can consider the continuum limit of the Hubbard model—first write $\ell = La_0$, i.e., split the chain into L segments of lattice spacing a_0. In this case the position x is na_0. The continuum field operator in terms of the real and wavenumber space operators is

$$\psi_a(x) = \lim_{a\to\infty} \frac{c^\dagger_{n,a}}{\sqrt{a_0}} = \frac{1}{\sqrt{\ell}} \sum_{k=-\infty}^{\infty} e^{-iq_k x} c^\dagger_{k,a}$$

where $c_k = 2\pi k/\ell$. The inverse is then

$$c^\dagger_{k,a} = \frac{1}{\sqrt{\ell}} = \int_0^\ell dx e^{iq_k x} \psi^\dagger_a(x)$$

Bearing in mind that the continuum limit takes the sum to an integral as

$$\sum_{j=1}^{L} a_0 f(na_0) \xrightarrow{a_0\to\infty} \int_0^1 dx f(x)$$

(of course we could shift $(0,\ell) \to (-\ell/2, \ell/2)$ and $\ell \to \infty$ for the infinite system limit). The kinetic term becomes

$$\begin{aligned} -t\sum_{j=1}^{L} \left(c^\dagger_{j,a} c_{j+1,a} + c^\dagger_{j+1,a} c_{j,a}\right) &= -t \sum_j a_0 \left(\frac{1}{a_0} c^\dagger_{j,a} c_{j+1,a} + \frac{1}{a_0} c^\dagger_{j+1,a} c_{j,a}\right) \\ &= -t \int dx (\psi^\dagger_a \psi_a(x+a_0) + \psi_a(x+a_0)^\dagger \psi_a(x)) \end{aligned}$$

Now, recall that translations are generated by

$$\psi(x+a_0) = e^{a_0\partial_x}\psi(x) = \psi(x) + a_0\partial_x\psi(x) + \frac{1}{2}a_0^2\partial_x^2\psi(x) + \cdots$$

Consequently, we can write

$$\begin{aligned} &\psi_a(x)^\dagger \psi_a(x+a_0) + \psi_a(x+a_0)^\dagger \psi_a(x) \\ &\quad = \psi^\dagger_a \left(\psi_a + a_0\partial_x\psi_a + \frac{1}{2}a_0^2\partial_x^2\psi_a\right) + \left(\psi^\dagger_a + a_0\partial_x\psi^\dagger_a + \frac{1}{2}a_0^2\partial_x^2\psi^\dagger\right)\psi_a \\ &\quad = 2\psi^\dagger_a\psi_a + a_0\left(\psi^\dagger_a\partial_x\psi_a + \partial_x\psi^\dagger_a\psi_a\right) + \frac{a_0^2}{2}\left(\psi^\dagger_a\partial_x^2\psi_a + \partial_x^2\partial^\dagger_a\psi_a\right) \\ &\quad = 2\psi^\dagger_a\psi_a + a_0\partial_x(\psi^\dagger_a\psi_a) + \frac{a_0^2}{2}\left(\partial_x^2(\psi^\dagger_a\psi_a) - 2(\partial_x\psi^\dagger_a)(\partial_x\psi_a)\right) \end{aligned}$$

If we assume the fields aren't supported at the boundaries, the total derivative terms vanish, and we're left with

$$\int dx \left[-2t\psi^\dagger_a\psi_a + ta_0^2(\partial_x\psi_a)^2\right]$$

Now, the first term is an IR divergence which ultimately may be ignored. The interaction term is

$$\begin{aligned} a_0 u(1-2n_{j\uparrow})(1-2n_{j\downarrow}) &= a_0 \pm u(1+2n_{j\uparrow}-2n_{j\downarrow}+4n_{j\uparrow}n_{j\downarrow}) \\ &= \pm u\left(a_0 - 2\psi_a^\dagger\psi_a + \frac{2}{a_0}\psi_a^\dagger\psi_b^\dagger\psi_a\psi_b\right) \end{aligned}$$

The sign ambiguity comes from different continuum limits corresponding to the Shiba transformation of the $c_{j,a}$ operators. In aggregate, the continuum Hamiltonian is

$$H = \int dx\, \mathcal{H}(x)$$

with Lagrangian density

$$\mathcal{H}(x) = -2(t \pm u)\psi_a^\dagger\psi_a + a_0^2(\psi_x\psi_a^\dagger)^2 \pm \frac{2u}{a_0}\psi_a^\dagger\psi_b^\dagger\psi_b\psi_a$$

The interaction term is manifestly UV divergent due to the inverse a_0 and will need to be regularized. Ignoring the IR divergent first term, we we find that the continuum tight-binding model recovers the nonlinear Schrödinger equation. Also, observe that the absence of a time derivative, meaning that this theory's Lagrangian actually coincides with its Hamiltonian (up to a sign).

4.3.1 Low Energy EFT

We have computed the continuum Hamiltonian of the Hubbard model—now, we expand on this result to find the low energy effective field theory of the model. We find that at less than half-filling the model is described by the $\mathrm{SU}(2)_1$ WZW model, and at half-filling it is described by the SU(2) Thirring model.

4.4 LSM and 't Hooft Anomalies ...

4.5 General effective actions 4/13/24

4.5.1 SSB and nonlinear actions

Spontaneous symmetry breaking takes us from an internal symmetry group (which here I assume to be a Lie group)G to a subgroup H by "breaking" the generators of a coset set G/H[1]. The theory should remain G-invarint, but while the vacuum expectation value

[1]Coset set = set of cosets gH. People use the term coset space but I don't think G/H is a vector space.

or order parameter is obviously invariant under H, the G/H fluctuations about the VEV, parameterized by Goldstone bosons, will transform nonlinearly (i.e., not as a linear representation). In fact, as is frequently drilled into students' heads, G/H is not even guaranteed to be a group, as H need not be normal, but this is not essential. To get a handle on these transformation rules, it is useful to think of the space of field configurations as a spacetime-independent manifold $\mathcal{M}$ on which G acts—for instance, a U(1) theory corresponds to the manifold S^1. $\mathcal{M}$ then has coordinates ϕ^i, at least locally. This provides us a dictionary between the actions of the internal symmetry on the vacuum field configuration and the more concrete theory of group actions.

For instance, taking the field to condense to ϕ_0, the stabilizer of the vacuum corresponds to H_{ϕ_0}, the unbroken part of the symmetry group (we endow H with a subscript ϕ_0 as, strictly speaking, what portion of G is unbroken depends on what the field condenses to). Moreover, the orbit of ϕ_0 under G traces out a submanifold of $\mathcal{M}$ which is in fact equivalent—let's say isometric—to G/H. This is because the coset space comprises sets gH_{ϕ_0} for $g \in G$, from which we can see that the uniue cosets are those that correspond to elements of G that don't leave ϕ_0 invariant (are not in H_{ϕ_0}). We will strictly think of G/H as a manifold with points denoted by ϕ unless otherwise noted.

Now, let us choose a set of coordinates ϕ^a of dimension $\dim G/H_{\phi_0}$ and χ^i of dimension $\dim H_{\phi_0}$ that paramterize G such that $(\pi^a, 0)$ traces out G/H_{ϕ_0} as a submanifold. We asume H_{ϕ_0} acts linearly on χ^i (we can always construct π and χ such that this is true). The entire theory should continue to be G invariant—this includes $(\pi^a, 0)$, of course, which implies that the action of H_{ϕ_0} on (π^a, χ^i) possesses an invariant subspace in the sense that π^a and χ^i are acted on independently by elements of the stabilizer. This just guarantees us that the H_{ϕ_0} transformation rules don't mix π and χ coordinates.

Now, we can reach any $\phi \in G/H_{\phi_0}$ from ϕ_0 by any element of G that is in that coset ϕ; pick one such element of that coset to be the coset representative, denoted $U(\pi)$ to remind us that this point in G/H_{ϕ_0} as a manifold is given by some configuration of π^a fields. $U(\pi)$ is notably an element of G, so U can be thought of as a map from a Goldstone boson field configuration π to the full group G with image G/H_{ϕ_0}. Perhaps it is most concrete to think of $U(\pi)$ as the exponential map $U(\pi) = \exp(i\pi^a Q_a)$ for generators Q of the Lie algebra $\mathfrak{g}/\mathfrak{h}$. Under this view $U(\pi)$ transforms under elements of H_{ϕ_0} as

$$U(\pi) \to U(\pi') = hU(\pi)h^{-1}$$

where, crucially, π' is the field configuration obtained by acting on π with the linear representation of h. While $U(\pi) \to gU(\pi)g^{-1}$ for any $g \in G$, we cannot so easily write $gU(\pi)g^{-1} = U(\pi)$ for some π' related linearly to π. One obtains $U(\pi') = hU(\pi)h^{-1}$ by noting that the action of $U(\pi)$ on the vacuum ϕ_0 followed by h is the same as applying h to ϕ_0 to shift to a new coordinate system and then applying $U(\pi')$ where π' is π in the h-shifted frame (think of changes of bases in linear algebra). That is,

$$hU(\pi)\phi_0 = U(\pi')h\phi_0 \implies hU(\pi)h^{-1} = U(\pi')$$

How does this look for a general $g \in G$? Well, we can factorize $g = kh$ where $k \in G/H$ (pedantic point but I think there's a theorem somewhere that says we can always do this, I just worry that G/H needs to be a subgroup), thus

$$gU(\pi) = khU(\pi) = kU(\pi')h$$

But $kU(\pi')$ is itself another coset representative which we may write $U(\pi'')$, but we will just relabel it to $U(\pi')$ to be concise. This

$$gU(\pi) = U(\pi')h(g,\pi)$$

were we have written $h(g,\pi)$ to emphasize that the particular h will depend on what the coset representative and g are. The intuitive way to view this is as the statement that we have $\pi \to \pi'$ under G/H ($U(\pi')$ being the representative for the new coset), but we additionally need the H transformation, i.e., we need an $h \in H$ to ensure that this is the correct coset representative for the action of g. $U(\pi')h(g,h) \in G/H \times H \simeq G$, more or less. Thus, bearing in mind that all of this is ultimately G-invariant, we now have the additional knowledge of how roguhly the Goldstone bosons transform:

$$U(\pi) \to U(\pi') = gU(\pi)h(\pi,g), \qquad \chi^i \to \chi'^i = D^{(\chi)}(h(\pi,g))^i_j \chi^j$$

for a representation $D^{(\chi)}$ of G. The left equation above is what is meant by a nonlinear realization of the symmetry transformation. This is admittedly abstract (how do we find $h(\pi,g)$, for instance?) but it is possible to get finer details out of this construction; see Brauner §7.3.1. This is nonetheless not essential to the present discussion and remains an entry for another day.

4.5.2 General effective actions

We want the effective action to be fully G invariant, but the action transforms nonlinearly under G/H transformations—that is, under a variation of the Goldstone bosons. We would like to show that those transformations are then G invariant, so that indeed the action is fully G invariant.

We will show that the allowable terms are classified by $H^{d+1}(G/H, \mathbb{R})$.

References

- D'Hoker, Weinberg, "General Effective Actions", arXiv:hep-ph/9409402.
- Brauner, *Effective Field Theory for Spontaneously Broken Symmetry* (2024). See Ch. 7 in particular.

4.6 CFT Miscellany ...

Schottenholer: scaling dimensions, central charge, structure constants.

4.7 Toy example of renormalization ...

4.8 Generalized Gibbs ensemble ...

4.9 Mean field ansatze of spin liquids ...

Parton = slave-boson is

$$\mathbf{S}_i = f^\dagger_{\alpha i} \boldsymbol{\sigma}_{\alpha\beta} f_{\beta i}$$