2019 physics diary / problem list

Ethan Lake

June 26, 2019

Preface: Disclaimer: I only know that my answers are correct for $\sim 2/3$ of the problems, and of course even correct answers are bound to have a nonzero number of typos / mistakes.

A few problems are listed in multiple section headings as appropriate.

May 23 — Anomalous dimension and correlators from the RG approach

Today we'll (try) to explain exactly what people mean by diffeomorphism invariance in GR.

Solution:

The nth order 1PI vertex is defined trough the effective action Γ:

$$\Gamma[\phi] = \sum_l \frac{1}{l!} \int \prod_{i=1}^{l} d^d x_i \, \phi(x_i) \Gamma^{(l)}(x_1 \ldots, x_l). \tag{1}$$

Since the effective action is dimensionless, we see that using the scaling of ϕ around the Gaussian fixed point, the nth order 1PI diagram has dimension

$$[\Gamma^{(n)}] = nd - \frac{n(d-2)}{2}. \tag{2}$$

May 25 — Hermitian conjugation and inner products for Euclidean CFTs

Today we'll be looking at , and will essentially be doing a collection of several exercises in David Simmons-Duffin's class notes on CFT.

Solution:

We map operators to Euclidean signature via

$$\mathcal{O}^l_E(\tau, \mathbf{x}) = (-i)^{\delta_{l,0}} \mathcal{O}^l_L(t, \mathbf{x}). \tag{3}$$

The reason for the prefactor is because it is required to ensure that $O(d)$ or $O(d-1,1)$ transformations commute with the processes of continuing between signatures. We can see that this prefactor is correct by mapping a spin-1 operator to real time, doing a Lorentz tform, and then mapping back to E time:

$$\mathcal{O}^l_E(\tau, \mathbf{x}) \mapsto (-i)^{\delta_{l,0}} \mathcal{O}^l_L(t, \mathbf{x}) \mapsto \Lambda^l_m (-i)^{\delta m,0} \mathcal{O}^m_L(t, \mathbf{x}) \mapsto i^{\delta_{l,0} - \delta_{m,0}} \mathcal{O}^m_E(\tau, \mathbf{x}). \tag{4}$$

If this is to be a sensible way of going between E and L signatures, then an $O(d)$ rotation R^{lm} must be related to the $O(d-1,1)$ rotation Λ^l_m via (our use of is here is legit because we are considering the complexification of the orthogonal groups)

$$R^{lm} = \Lambda^l_m i^{\delta_{l,0} - \delta_{m,0}}. \tag{5}$$

Let's check to see that this works: if it works, then the LHS needs to preserve δ_{lm}. Being slightly callous about index placement, and working in mostly positive signature, we have

$$R^{lm} R^{ln} = \Lambda^l_m \Lambda^l_n i^{2\delta_{l,0} - \delta_{m,0} - \delta_{n,0}} = \Lambda^l_m \Lambda_{ln} i^{-\delta_{m,0} - \delta_{n,0}} = \eta_{mn} i^{-\delta_{m,0} - \delta_{n,0}} = \delta_{mn}, \tag{6}$$

and so the RHS of (5) is indeed orthogonal. A general spin operator then maps as

$$\mathcal{O}^{l_1 \ldots l_n}_E(\tau, \mathbf{x}) \mapsto (-i)^{\sum_i \delta_{l_i,0}} \mathcal{O}^{l_1 \ldots l_n}_L(t, \mathbf{x}). \tag{7}$$

Now in Euclidean signature there is no natural choice of inner product. Since we want to ground the whole formalism in real time where physics is defined, we then use the inner product in real time to construct one in Euclidean time. The Euclidean Hermitian conjugation must then reverse the Euclidean time τ—different choices of what τ is lead to different

Hilbert spaces, and hence to different quantizations of the theory. In addition to reversing τ, conjugation also changes the signs of vector indices with time components, since these have an extra i associated to them in Euclidean signature, as described above. Therefore the Euclidean Hermitian conjugation acts as

$$[\mathcal{O}_E^{l_1\cdots l_n}(\tau,\mathbf{x})]^\dagger = (-1)^{\sum_i \delta_{l_i,0}}\mathcal{O}_E^{l_1\cdots l_n}(-\tau,\mathbf{x}). \tag{8}$$

This is actually very sensible: since † flips the τ coordinate, vector indices pointing in the τ direction get minus signs.[1]

When we quantize on the cylinder $\mathbb{R}\times S^{d-1}$, τ is just the z-coordinate of the cylinder, and we can use the above formula for Hermitian conjugation. However, since we will more often work in radial quantization on the plane, we need to port this definition of Hermitian conjugation into radial quantization conventions. The zero of Euclidean time is the unit sphere in $\mathbb{R}^d$, and so $\tau\mapsto-\tau$ corresponds to the inversion $x^\mu\mapsto x^\mu/x^2$. The minus sign in the mapping (8) is accounted for in radial quantization by contracting all the indices of $\mathcal{O}_E$ with the matrix $I^\mu_\nu(x)=\delta^\mu_\nu-2x^\mu x_\nu/x^2$, taking $x^\mu\mapsto -x^\mu$. Finally, we need to remember that the cylinder and the plane have metrics differing by a Weyl rescaling. Since $r\to e^\tau$, the Weyl rescaling is determined via $r^2(dr^2/r^2+d\Omega^2_{d-1})\to e^{2\tau}(d\tau^2+d\Omega^2_{d-1})$, so that the rescaling factor is $e^{2\tau}$. Now we need to remind ourselves that if we have two different coordinate systems with metrics such that $ds_1^2=e^{2w}ds_2^2$, then an operator with scaling dimension Δ is given in the two coordinate systems by $\mathcal{O}(x_1)=e^{-w\Delta}\mathcal{O}(x_2)$. This means that $\mathcal{O}_E(x^\mu)=e^{-\tau\Delta}\mathcal{O}_E(\tau,\mathbf{x})$. Therefore we can determine how † acts on scalar operators in radial quantization via

$$\mathcal{O}(x^\mu)^{\dagger_r}=e^{-\Delta\tau}\mathcal{O}(\tau,\mathbf{x})^\dagger=e^{-\Delta\tau}\mathcal{O}(-\tau,\mathbf{x})=e^{-\Delta\tau}(e^{\Delta(-\tau)}\mathcal{O}(x^\mu/x^2))=r^{-2\Delta}\mathcal{O}(x^\mu/x^2). \tag{9}$$

For operators with spin we just add in the I matrices we mentioned above to this equation, and hence

$$[\mathcal{O}_E^{l_1\cdots l_n}(x^\mu)]^{\dagger_r}=r^{-2\Delta}I^{l_1}_{m_1}(x)\cdots I^{l_n}_{m_n}(x)\mathcal{O}_E^{m_1\cdots m_n}(x^\mu/x^2). \tag{10}$$

The factor of $r^{-2\Delta}$ is crucial here, since it is needed to make inner products like

$$\langle\mathcal{O}|\mathcal{O}\rangle=\langle 0|\mathcal{O}(0)^{\dagger_r}\mathcal{O}(0)|0\rangle=\lim_{r\to\infty}r^{2\Delta}\langle 0|\mathcal{O}(r)\mathcal{O}(0)|0\rangle \tag{11}$$

finite in radial quantization (and properly dimensionless).

May 27 — Cuprate phenomenology review

The students in Senthil's strongly correlated electrons class this semester had to do final mini-projects; mine was on cuprate phenomenology. I'm using it as a diary entry today just so I can keep it around, for posterity's sake.

[1] An easy mistake to make is to apply this formula blindly to everything with a vector index. If do this, we would conclude that e.g. $P_1^\dagger=-P_1$ and $P_l^\dagger=P_l$ for $l>1$ (here P_i is the physical momentum operator on e.g. the cylinder; we are not (yet) working in radial quantization, where of course the momentum is conjugated as $P^\dagger=K$). However, it's this is distressing, since we know that P_1 is the Hamiltonian, which should be Hermitian. In fact, the transformation rule is the opposite: P_1 is Hermitian, and the rest are anti-Hermitian (which can be seen e.g. by taking the † of $[P_1,\phi(0)]=\partial_\tau\phi(0)$, with $\phi(0)^\dagger=\phi(0)$). Likewise, one can show that $M_{lm}^\dagger=-M_{lm}$ for $l,m>1$, while $M_{1l}^\dagger=M_{1l}$. So, one should only apply our Hermitian conjugation rule to primaries.

In what follows, we will provide a lightning review of hole-doped cuprate phenomenology. The focus is on facts from experiments rather than on trying to explain things with toy models and numerics. Since the cuprates have been studied to death, we will be very selective, detailing only what we think are the most representative experimental findings. Our strategy will be to provide brief vignettes on several different experimental results, rather than going into detail about any one particular technique.

We'll skip the usual spiel about how "ever since their discovery, cuprates have…", and will instead just start by outlining our plan of action for this mini-review. Since ~ 4 pages is hilariously little to do justice to a field with $O(10^5)$ research articles and a 30+ year history, and since the theoretical situation is often rather confused, we will refrain from going into undue detail about the theoretical side of things; the focus will be on experiments. In particular, we will emphasize older / more established experimental results—they already offer a wealth of information that theorists have yet to explain.

We'll concentrate on hole-doped cuprates. We will be particularly concerned with the under-doped (UD) side of the phase diagram since it contains the pseudogap (PG), which is a phenomenon endemic to the high temperature superconductors (HTSCs) and strongly tied to the SCing phase—the strange metal behavior seen on the over-doped (OD) side occurs in other materials besides HTSCs (e.g. heavy fermion metals), and so we will devote less attention to it. All the essential information needed to achieve a theoretical understanding of the HTSCs may well already be out there—all that is needed is for theorists to meditate and reflect on the available experimental results. With that in mind, let's get to the data.

7 Experiments

The bulk of this mini-review consists of brief snapshots of a few different experimental strategies adapted for studying the cuprates, listed roughly in the order in which they began to be fruitfully applied to the HTSC problem.

NMR

NMR has been used in several contexts for studying HTSCs. Since it can probe the magnetic environment that each distinct element in a given cuprate compound lives in, it can determine precisely where the holes that contribute to the spin interactions are located. For example, consider $YBa_2Cu_4O_8$, which is doped by changing the O concentration in the layers between the CuO planes. In principle the doped holes could be located in O orbitals that are unrelated to the holes in the Cu sites. However, it was found[**?**] that (for UD samples), the Cu and O Knight shifts have the same (unusual; more on this in a sec) T-dependence. This tells us that the spin susceptibility in fact comes from a single band involving strongly hybridized Cu-O orbitals, the knowledge of which is important for setting up theoretical models for the HTSCs.

More prominently, NMR was the first tool used to identify the PG present in the UD side of the phase diagram. Figure 1 shows Cu spin relaxation rates in UD YBCO (chosen for NMR since it is the most ordered of the cuprates, which ensures that all the Cu atoms live in essentially the same magnetic field). For this material, $T_C \approx 80$ K. Below T_C the

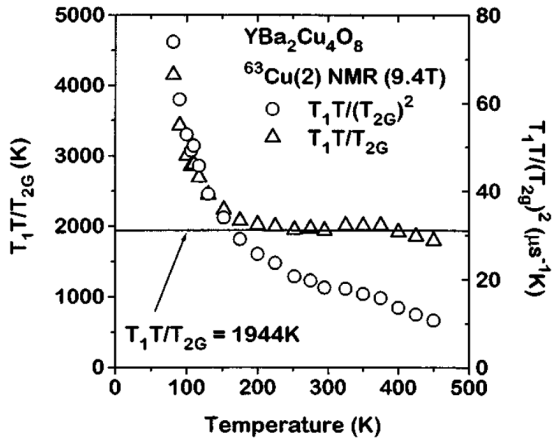


Figure 1: NMR relaxation rate data for the Cu line in UD YBCO, which has $T_C \sim 80$ K. Here the plot is the ratio of longitudinal relaxation time T_1 for the Cu nuclei (relaxation along the magnetizing field direction—depending on χ''; dominated by AF fluctuations) and the transverse relaxation time T_{2G} (relaxation in the plane normal to the magnetizing field—depending on χ') times the temperature T ($1/T_1$ depends linearly on T in typical metals). Taken from Ref.[**?**]

electron spins are locked into singlets and so they cannot provide a channel for the nuclear spins to relax back to the magnetizing field after being perturbed by the NMR pulse; hence the diverging relaxation time below 80 K. However, the behavior at $T \lesssim T_C$ isn't what we expect from a BCS SC, for two reasons: 1) the nuclear relaxation rate falls off only as a power law below T_C, not as an exponential, implying that there are still some spin degrees of freedom active below T_C, so that the gapping is not complete, and 2) the absence of a sharp decrease in the relaxation rate just below T_C, which occurs in BCS theory because of the added spectral weight right above the gap.

The plot shows something else unusual: the loss in spin susceptibility with decreasing T happens well before T_C is reached (typical UD cuprates see a decrease in their spin susceptibilities by a factor of $\sim$ 3-5 between room temperature and T_C), and indeed on the plot, nothing special happens at 80 K. Instead, the scaling changes at $T^* \sim 200$ K$\gg T_C$. Therefore some nonzero fraction of the spins (but not all of them) must be getting gapped and locked away into singlets well before SC occurs—hence the reason why this region is given the moniker "pseudogap". In what follows, we will continue to use T^* to refer to the temperature marking the upper boundary of the PG—different experimental probes see different manifestations of the PG phenomenon, but for the most part all agree on a common temperature at which it starts to occur.

Photoemission

ARPES is the method of choice for studying the band structure of the cuprates—by bombarding the sample with photons incident on the ab plane and measuring the kinetic energy of the ejected electrons at various ejection angles, one can build up a comprehensive map of the density of states below the Fermi level. In the SCing state, it is generically found that

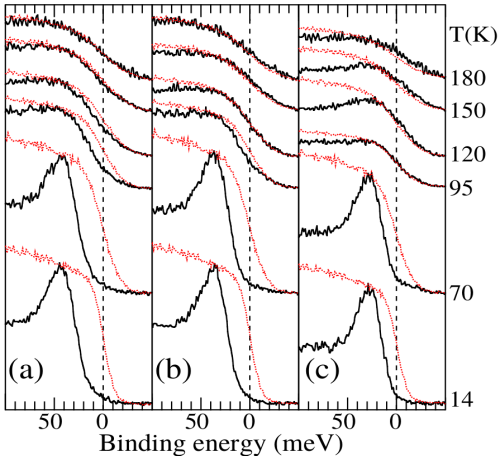


Figure 2: One of the cooler ARPES plots. The black curves are photoemission spectra from lightly UD BSCCO ($T_C = 85$ K), and the red curves are from a sample of platinum attached to the sample to get a reference for the zero on the energy axis. The leftmost panel is taken at an antinode $\sim (\pi, \pi/5)$, the rightmost is taken closer to the Fermi arc location at $\sim (2\pi/3, \pi/5)$, and the middle panel is taken between the two. Taken from Ref.[**?**]

gapless points survive along the nodal lines $(\pm\pi/2, \pm\pi/2)$, with gapped antinodes spaced in between; this helps confirm the nature of the d-wave pairing symmetry. Above T_C a gap is found to remain at the antinodes for UD samples, with Fermi arcs at the node locations—this is the PG regime, which persists up to essentially the same temperature T^* as seen in NMR studies.

Photoemission spectra for UD $Bi_2Sr_2CaCu_2O_8$ (which is well-suited for ARPES studies since it cleaves easily; the energies employed in typical experiments mean that the photons don't penetrate far into the bulk, so a very clean surface is needed) is shown in Figure 2. The leftmost panel shows spectra (black) for various T, taken at k_F for an antinodal point in the BZ. This is where the PG is biggest—the midpoint of the sample's leading edge doesn't meet zero energy until the hottest curve, so that we can identify $T^* \sim 180$ K. By looking for the value of T at which the coherent peak disappears, we identify the critical temperature as $T_C = 85$ K. Note that the disappearance of the peak means that not only Cooper pairs loose their coherence above T_C, but that single electrons loose their coherence as well. The middle and right panels show spectra from closer to the node, where both the SCing gap and the PG are seen to be suppressed.

More generally, ARPES studies on UD samples show that the momentum dependence of the PG tracks that of the SCing gap and evolves smoothly into it as T is lowered. This can be taken as evidence that the PG is characterized by the existence of bound fermion pairs which lack phase coherence. This is true within a finite temperature window above T_C, but not likely true all the way up to T^*, as we will see. Interestingly, the size of the SCing gap is found to increase with underdoping, even though T_C decreases. ARPES has also been used to map out the OD side of the phase diagram, where metallic states with large FSs are generically found.

Transport

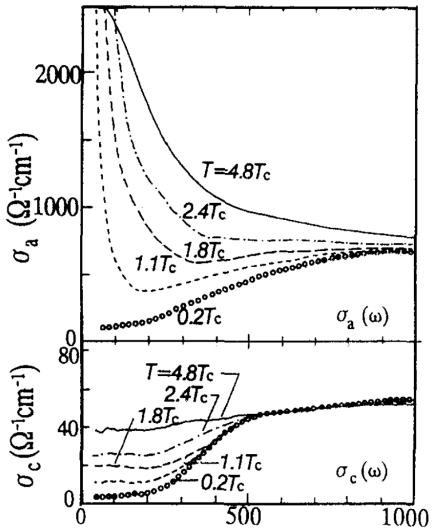


Figure 3: Conductivity in-plane (upper) and out-of-plane (lower) as a function of ω (in cm^{-1}) for UD YBCO ($T_C \sim 80$ K). Several optical phonon lines have been removed in the lower plot to display the electronic background more clearly. Taken from Ref.[**?**]

Optical conductivity stands out among the different transport-related experiments in the wide variety of information it offers. Figure 3 shows optical conductivity measurements for UD YBCO, chosen because it is the cleanest of the commonly studied cuprates (so that it has nice shiny surfaces). The upper panel shows the ab-plane conductivity: above T_C we see a standard Drude peak that starts at ~ 500 cm^{-1} and narrows for lower temperatures. At intermediate temperatures above T_C we see a depression around 400 cm^{-1}; in OD samples this feature is absent, suggesting that it is related to the PG (which NMR tells us occurs at $T^* \sim 3T_C \sim 250$ K).

The lower panel shows the optical conductivity along the c-axis. Overall the conductance is low and flat, telling us that electron hopping in this direction is largely incoherent. There is not even remotely a sign of a Drude peak like the one seen in the in-plane data. Evidently there is more to the difference between the two conductivities than just the fact that the material is anisotropic.

The c-axis conductivity shows very clear signs of the PG: at low frequencies there is a depressed gap starting at T^*, with no qualitative change as T_C is crossed. Consistent with ARPES measurements, the gap is frequency-independent with changing T. The plots for near-optimally-doped samples in the SCing state turn out to look similar to the ones for UD samples in the normal state, telling us that the PG and the SCing gap are likely intimately connected. Also note that taken together, the conductance data jives with the RVB idea: transport out of the plane is suppressed since it is only possible if singlets are broken (which necessitates paying a nonzero energy cost), while in-plane transport looks more typical, given that it is facilitated by the gapless holes.

One word of caution here is that different cuprates have very different transport properties vis-a-vis the magnitudes and ratios of the in-plane and out-of-plane conductivities. For

example, the c-axis conductivities for YBCO and BSCCO differ[**?**] by a factor of $\sim 4 \times 10^4$! Even more distressingly, the general behavior of the T dependence of the normal-state c-axis transport seems to vary appreciably across the different cuprates—as far as we're aware, the reason for this, and its possible significance to the problem at hand, is still not well-understood.

STM

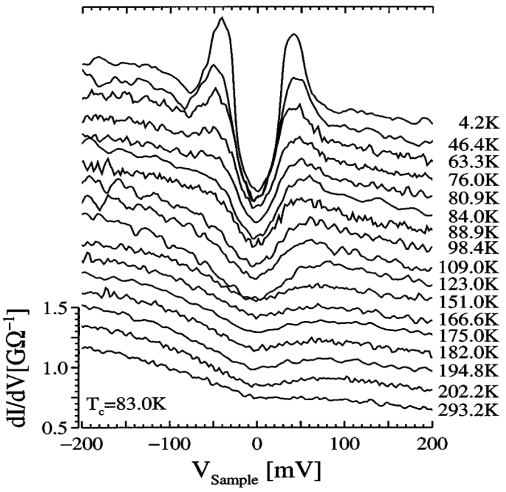


Figure 4: dI/dV (alias local DOS) as a function of sample bias in an STM measurement on UD BSCCO. Note that the width of the gap is unchanged above $T_C = 83$ K—the only special feature at T_C is the disappearance of the coherent peaks on either side of the gap. Taken from Ref.[**?**]

STM techniques, like ARPES studies, allow a glimpse into what the electrons near the Fermi level are doing. Figure 4 shows tunneling conductance data for UD BSCCO (ideal for STM because it cleaves easily—STM needs good surface quality since the electrons in HTSCs have short coherence times and hence STM only sees the first few layers of a material). This material has $T_C = 83$ K, and we see temperatures below this are accompanied by broad peaks separated by $2\Delta \approx 90$ meV. As far as we're aware, the exact interpretation of the incoherent features beyond the peak, particularly the dips just below the peak (which are corroborated by ARPES studies) is still rather controversial.

The PG manifests itself as the gap in the DOS at zero bias, which occurs smoothly out of the SCing gap and persists to well above T_C. Going against BCS intuition, the gap is pretty much frequency independent across the entire temperature range, and when doping is increased, the gap decreases (even though T_C increases).

STM can do more than just confirm ARPES results, though—it has also been used to reveal the significantly heterogeneous nature of the SCing state in the UD region. Figure 5 shows tunneling conductance curves as a function of spatial position in a sample of UD BSCCO, scanning a length of 140 Å. Along the path of the STM scan, we see SCing regions loose their peaks and turn into PG-like regions (while maintaining the size of the gap). These results indicate the presence of inhomogeneities in the SCing state on the scale of $O(10)$ Å.

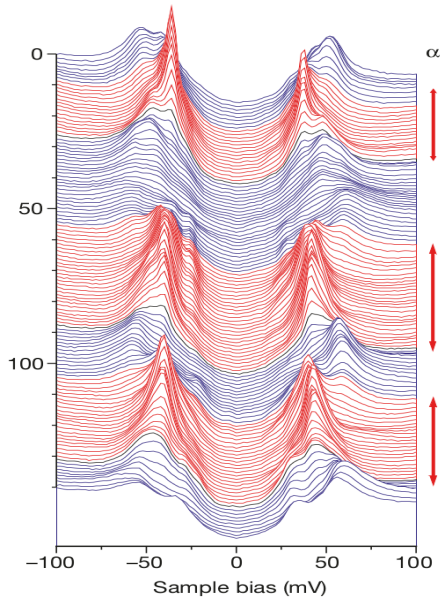


Figure 5: Very pretty dI/dV curves as a function of sample bias and spatial position (in Å) along a 140 Å path in a sample of UD BSCCO. The red areas are SCing; the blue areas have PG-y features. Taken from Ref.[**?**]

Like the pseudogap itself, these inhomogeneities are seen to become larger with underdoping; for severely UD samples they persist to very low temperatures, indicating that these regions may be some zero-temperature manifestation of the PG (see e.g. Ref[**?**]). Also note that if one restricts to low ($\sim$ 10 meV) energies, the density of states looks homogeneous—this suggests that the low energy modes are spatially coherent, while the higher energy modes are not.

STM can also be used to look for things like CDW order and can be used to peer into the states inside of vortex cores; for space reasons we will sadly not get to discuss these applications.

Magnetic fields

The fate of superconducting fluctuations near T_C can be studied by utilizing the Nernst effect: in the presence of a thermal gradient, magnetic field, and nonzero pairing amplitude, the entropy-carrying magnetic vortices will flow along $-\nabla T$, producing a transverse electric response due to the "phase flux" carried by the vortices. Figure 6 shows the magnitude of the Nernst effect (the strength of the electric response) at different points in the phase diagram of LSCO. From the extent of the Nerst region, we can infer a residual nonzero pairing amplitude for a large range of temperatures above T_C, implying that SC is killed at T_C by lack of phase coherence and the establishment of a vortex liquid, and not by the breaking-apart of pairs. Another important point is the fact that profile of the Nernst region has a shape similar to the T_C dome (in contrast to T^*), and that the Nernst region stops before T^* is reached—therefore residual SCing phase fluctuations cannot be the sole explanation for the PG.

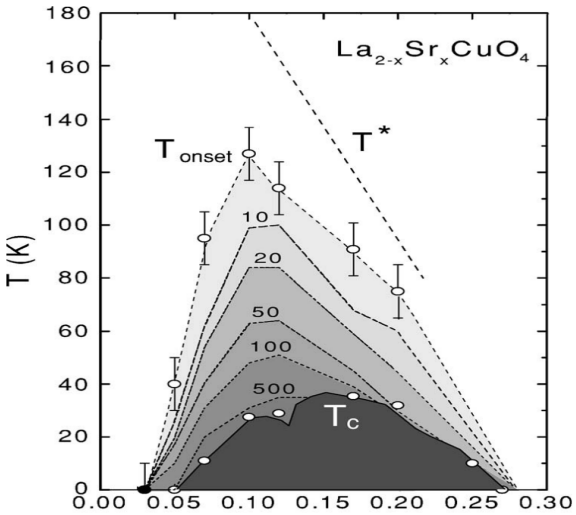


Figure 6: Nernst effect in the phase diagram (T vs x) of $La_{2-x}Sr_xCuO_4$; the contours indicate the value of the Nernst coefficient. The important thing is that the Nersnt region does not extend up to T^*. Taken from Ref.[**?**]

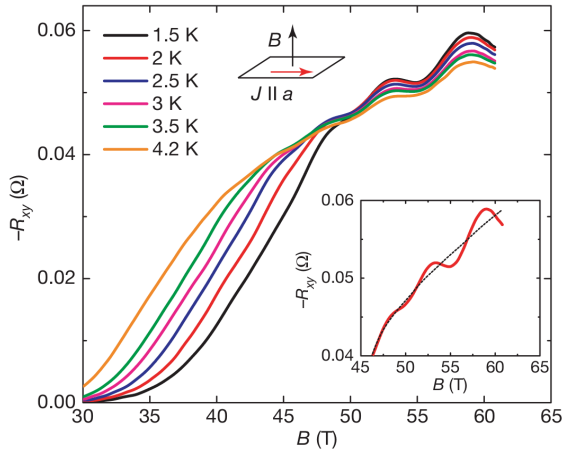


Figure 7: Hall resistance in UD YBCO at several different temperatures, showing quantum oscillations. After subtracting out the linear part, one finds a frequency for the $T = 2$ K curve of $f \approx 500$ (in Tesla). Taken from Ref.[**?**]

The measurement of quantum oscillations is a relatively contemporary experimental probe, but has offered a lot of interesting information about the potential phases "hiding" behind the SC dome. In Figure 7 we show quantum oscillations observed in the Hall resistivity for a sample of UD YBCO (chosen because YBCO more easily yields cleaner+more ordered crystals relative to the other cuprates; quantum oscillations are very sensitive to disorder). These measurements are taken well below the upper critical field H_{c2}, which is likely a factor of ~ 3 larger than the field at which the resistive transition out of a SCing state happens.

One things that immediately jumps out is the fact that the Hall coefficient is *negative*, indicating that the charge carriers have negative charge—to zeroth order this is unexpected, on account of the material being doped with holes, not electrons. Since the period of the oscillations tells us the in-plane FS area, one can conclude that in the UD sample shown, the FS occupies a measly $\sim 2\%$ of the BZ, implying the existence of small electron pockets.

Another somewhat surprising result is the fact that we see quantum oscillations at all: this implies the existence of a FS, but we know that the PG state doesn't have one (it has open Fermi arcs, which can't give closed orbits to contribute to the oscillations). Therefore the magnetic field which destroys the SCing state doesn't "reveal" the underlying $T = 0$ PG state. Evidently either the PG doesn't survive down to $T = 0$, or the magnetic field does more than just killing superconductivity.

In contrast to the UD sample shown above, OD samples show a much higher oscillation frequency with positive R_{xy}, implying the existence of the large (hole-like, in this case) FS that tools like ARPES tell us to expect from that part of the phase diagram.[2] Therefore the outstanding question is how the small electron pockets evolve into the large hole-like FS, and whether the FSs seen at high fields are actually germane to the zero field problem, and not just a consequence of the existence of the applied field.

While these measurements can only tell us the area of the FS (and not its shape or position within the BZ), experience tells us that one likely scenario is for the UD side to contain small electron pocket(s) arising from some sort of density wave order[3], since translational symmetry breaking lets us get around the constraints of Luttinger's theorem, which we need to do in order to reconcile the UD and OD FS measurements. The presence of topological order is another way to circumvent Luttinger's theorem, and is at least a possibility in not too UD YBCO, where no ordered state has been seen.[**?**]

Questions

Many questions that I have are likely un-answerable in a definitive sense at this time. Some questions that I have yet to answer, but which I think are potentially answerable in a precise

[2]High frequency oscillations means closely spaced energy levels, which means large orbits and hence a large FS.

[3]For a typical example, consider SDW order at the AFM vector $Q = (\pi, \pi)$ occurring on top of a large circular hole-like FS centered on Q—this is the FS seen in the OD side of the phase diagram, as revealed by ARPES. Building up SDW order means identifying Q with 0 in the BZ, which has the effect of superimposing a copy of the FS centered at Q on top of the original FS. The places where the original FS and its double meet then pinch off, leaving behind a collection of electron pockets at $(0, \pi), (\pi, 0)$ and hole pockets at $(\pm\pi/2, \pm\pi/2)$.

way, are

- In theoretical treatments, is it a bit hasty to work with a 1-band model in 2 dimensions? For example, the Coulomb interaction out of the CuO planes is (presumably?) not drastically screened, and yet in many theoretical treatments it is ignored.

- Theoretical treatments usually ignore the heterogeneity caused by the way the cuprates are doped. To what extent is it easy to distinguish the heterogeneity from other effects? For example, when we see comparatively broad features in ARPES data, what part of this is due to interesting modifications to the self-energy, and what part is just due to the doping potential? Or, when we observe strong heterogeneity it STM experiments, how do we distinguish between effects caused by electron interactions + electronic phase separation and effects caused by disorder?

- When exactly should we have confidence in treating all the different cuprates as one material? As an example question, we have little STM data on LSCO, since it doens't cleave well—to what extent do we expect conclusions about e.g. tunneling conductance and spatial heterogeneity, obtained by doing STM on e.g. BSCCO, to hold for LSCO?

- (possibly un-answerable at present) How important are the observations of different orders (e.g. the spin+charge stripe order in LSCO) to the high T_C problem? Given that such orders occur in pretty different ways across different classes of cuprates, I'm unsure about to what extent they have anything universal to say.

- This was already alluded to above, but why exactly are we confident that the conclusions obtained in high magnetic fields from e.g. quantum oscillation experiments are relevant for the zero-field HTSC problem? *update: We aren't, really. We...*

- Assuming that the conclusions about the FS topology that the high-field results give us are indeed relevant for the zero-field problem, is it really reasonable that the SCing state should be entirely insensitive to the drastic re-arrangement of the FS that needs to occur as the doping is varied?

- Is there an argument for why, in multilayer materials like YBCO, T_C increases with increasing number of layers? What aspect of the three-dimensional physics (which have mostly ignored) could be responsible for this?

May 26 — Unitarity bounds on CFT scaling dimensions

Today we'll be looking at unitarity constraints on scaling dimensions, and will essentially be doing an elaboration on an exercise in David Simmons-Duffin's class notes on CFT.

Solution:

First we will prove a bound on the scaling dimension of all operators with nonzero spin, namely that if $\mathcal{O}_n$ is an operator with spin n, then

$$\Delta_{\mathcal{O}_n} \geq d - 2 + n. \tag{12}$$

In particular, for $n = 1$, this tells us that the minimal possible scaling dimension is $d-1$, which is the dimension that a conserved current has. In fact we will see that a spin-1 operator is a conserved current iff its scaling dimension saturates the bound: conserved currents have no anomalous dimensions, and any spin-1 operator with no anomalous dimension is a conserved current.

First we have to define what we mean by "an operator of spin n" in dimensions other than 3. In here and what follows, spin n will always refer to an operator which transforms in the n-index symmetric traceless tensor representation of $SO(d)$. While for $SO(3)$ these representations exhaust all irreps,[4] this is of course not true for $SO(d > 3)$, when irreps assigned to antisymmetric tensors become possible to construct. However, conserved currents and energy momentum tensors and the like are usually all associated to symmetric traceless representations, so we will restrict our attention to them in what follows.

The proof goes by considering the inner product between two descendants in the conformal multiplet of a spin-n conformal primary $\mathcal{O}$:

$$\left(P^\mu|\mathcal{O}^{\boldsymbol{\beta}}\rangle\right)^\dagger P_\nu|\mathcal{O}^{\boldsymbol{\alpha}}\rangle = \langle\mathcal{O}^{\boldsymbol{\beta}}|K_\nu P^\mu|\mathcal{O}^{\boldsymbol{\alpha}}\rangle. \tag{13}$$

Here the notation is such that $\boldsymbol{\alpha}, \boldsymbol{\beta}$ label multi-indices in accordance with the symmetric and traceless nature of $\mathcal{O}$, so that e.g. $\mathcal{O}^{\boldsymbol{\alpha}} = \mathcal{O}^{\alpha_{(1},\ldots,\alpha_{n)}}$ (note the symmetrization). Note that this is not the inner product between two identical vectors and hence is not necessarily positive. Also note that we are not assuming any contraction between the μ, ν indices and those carried by $\mathcal{O}$. Using the commutator

$$[K_\mu, P_\nu] = 2\delta_{\mu\nu}D - 2M_{\mu\nu}, \tag{14}$$

where $M_{\mu\nu}$ is the $SO(d)$ generator for the rotation parametrized by μ, ν in the spin-n representation, we have (since $\mathcal{O}$ is primary)

$$\left(P_\mu|\mathcal{O}^{\boldsymbol{\beta}}\rangle\right)^\dagger P_\nu|\mathcal{O}^{\boldsymbol{\alpha}}\rangle = 2\langle\mathcal{O}^{\boldsymbol{\beta}}|\delta_{\nu\mu} - M_{\mu\nu}|\mathcal{O}^{\boldsymbol{\alpha}}\rangle. \tag{15}$$

Now we use a clever trick: since $M_{\mu\nu} = -M_{\nu\mu}$, we can write

$$M_{\mu\nu} = \sum_{i<j=1}^{d} [A^{ij}]_{\mu\nu} M_{ij}, \tag{16}$$

where $[A^{ij}]_{\mu\nu} = \delta^i_\mu\delta^j_\nu - \delta^j_\mu\delta^i_\nu$ is the generator matrix of the rotation matrix parametrized by the tuple ij, represented in the vector representation. Let us normalize the states by

$$\langle\mathcal{O}^{\boldsymbol{\beta}}|\mathcal{O}^{\boldsymbol{\alpha}}\rangle = \delta^{\boldsymbol{\beta}\boldsymbol{\alpha}}. \tag{17}$$

[4]Since any tensor with a pair of antisymmetric indices can be reduced contracting with the invariant symbol ϵ_{ijk}. This reduces the number of indices in the tensor and lets us turn antisymmetric tensors into symmetric ones. This fails for larger $SO(d)$ since contracting with the invariant ϵ symbol does not decrease the number of indices (for $d > 4$ it increases the number of indices).

Therefore we may write

$$\frac{1}{2}\left(P_\mu|\mathcal{O}^{\boldsymbol{\beta}}\rangle\right)^\dagger P_\nu|\mathcal{O}^{\boldsymbol{\alpha}}\rangle = \delta_{\mu\nu}\delta^{\boldsymbol{\alpha\beta}}\Delta - \sum_a (A^a \otimes M^a)^{\boldsymbol{\alpha\beta}}_{\mu\nu}, \tag{18}$$

where a runs over the generators of $\mathfrak{so}(d)$.

Hence we have obtained the inner product as a matrix element of a matrix constructed from a $\otimes$ of a matrix in the F irrep and one in the S_n irrep, where F is the fundamental (vector) irrep and S_n is the symmetric, traceless, n-index tensor irrep. To deal with this expression, we should reduce this $\otimes$ to its constituent irreps. This happens by writing

$$\begin{aligned}\sum_a (A^a \otimes M^a)^{\boldsymbol{\alpha\beta}}_{\mu\nu} &= \frac{1}{2}\sum_a \left[(A^a \otimes \mathbf{1} + \mathbf{1} \otimes M^a)^2 - (A^a)^2 \otimes \mathbf{1} - \mathbf{1} \otimes (M^a)^2\right) \\ &= -\frac{1}{2}\left(C_2(F \otimes S_n) - C_2(F) \otimes \mathbf{1} - \mathbf{1} \otimes C_2(S_n)\right),\end{aligned} \tag{19}$$

since the generators for the $R_1 \otimes R_2$ rep are $T^a_{12} = T^a_1 \otimes \mathbf{1} + \mathbf{1} \otimes T^a_2$. Here the minus sign is because our generators are anti-Hermitian (e.g. A^a is real and antisymmetric), and since the quadratic Casimir is positive definite and is hence $\sum_a (T^a)^\dagger T^a = -\sum_a T^a T^a$ (e.g. $L^2 = -\partial_\mu \partial^\mu$ since ∂_μ is anti-Hermitian).

The quadratic Casimir[5] for the traceless symmetric nth-rank tensor representation of $SO(d)$ is allegedly

$$C_2(S_n) = n(n+d-2). \tag{20}$$

I tried for a stupidly long time to check this, but didn't quite figure it out. The thing I did figure out was the dimension of the representation S_n—to get C_2 we would then need the index, which I couldn't couldn't compute. The computation of the dimension is preserved for posterity's sake in the following footnote.[6]

[5] Just "Casimir" is imprecise terminology since for general groups there are usually many different Casimirs we can construct (involving kth powers of the generators—each k-index invariant symbol gives us such a Casimir).

[6] We want to figure out the dimension of the n-index symmetric tensor irrep of $SO(d)$. A tensor transforming under this irrep has n indices, each of which can take d values: therefore counting the number of such tensors amounts to the number of ways we can place n different objects (the indices of the tensor) in d different buckets. Only the total number of objects that gets placed in each bucket matters—the objects being placed in the buckets are all identical, on account of the symmetric property of the tensor meaning that all the indices can be freely exchanged. Recall from stat mech class that the best way to think about this is to consider counting the number of ways to arrange n objects and $d-1$ dividers between the different buckets. Since the dividers and objects are all identical, this number is $\binom{n+d-1}{n}$. Thus the number of symmetric tensors is $(n+d-1)!/(n!(d-1)!)$.

However, we need to take care of the traceless condition. This reads $T^{\mu\mu\nu_1\ldots\nu_{n-2}} = 0$, where by symmetry it doesn't matter where in the index structure the two μs are. Note that the remaining indices $\nu_1, \ldots, \nu_{n-2}$ are symmetric, and so the traceless condition gives us a number of constraints equal to the number of rank-$(n-2)$ symmetric tensors. By the above this is $\binom{n+d-3}{n-2}$. Therefore the dimension of S_n is

$$\dim[S_n] = \binom{n+d-1}{n} - \binom{n+d-3}{n-2}. \tag{21}$$

For $d = 3$ we get $2n+1$ as expected, while for e.g. $d = 4$ we get $(n+1)^2$. From the form of the quadratic Casimir, it's clear that the formula for the index $T(S_n)$ is going to be complicated.

Let's now go into a basis appropriate for the $\oplus$ decomposition of $F \otimes S_n$ into constituent irreps. In this basis, we have

$$(P|\mathcal{O}\rangle)^\dagger P|\mathcal{O}\rangle = 2\Delta \mathbf{1} - C_2(F) \otimes \mathbf{1} - \mathbf{1} \otimes C_2(S_n) + \bigoplus_{R \in F \otimes S_n} C_2(R), \tag{22}$$

where we are now thinking of the LHS as a matrix, rather than a matrix element. Now since the LHS is the matrix whose entries are the inner products of basis vectors for the first descendants of $\mathcal{O}$, it must be positive definite. The potential for it to have negative eigenvalues is contained within the choices for the different irreps in the $\oplus$ term: to get the strongest bound on Δ we should look for representations $R \in F \otimes S_n$ such that $C_2(R)$ is minimized. Since C_2 is larger for larger irreps, we should choose the smallest irrep appearing in the $\oplus$ decomposition. The smallest irrep is S_{n-1}, which we get from contracting the vector index of F with any of the indices of the symmetric tensors in S_n. Therefore we must have

$$2\Delta \geq (d-1) + n(n+d-2) - (n-1)(n-3+d) \implies \Delta \geq n + d - 2, \tag{23}$$

as claimed. Note that this bound only applies to primary operators. For example, one shouldn't worry that the dimension of A^μ in free Maxwell theory is $(d-2)/2 < d-1$, since A^μ is not a primary (it is not even an operator in the physical Hilbert space).

Now we will prove that all scalar operators have scaling dimension bounded from below by $\Delta \geq (d-2)/2$, which is saturated for the free scalar. Consider the inner product

$$0 \leq (P_\mu P^\mu|\mathcal{O}\rangle)^\dagger P_\mu P^\mu|\mathcal{O}\rangle = \langle \mathcal{O}|K_\mu K^\mu P_\nu P^\nu|\mathcal{O}\rangle. \tag{24}$$

Getting to the bound on Δ is now just a matter of algebra. Using the commutators

$$[D, K_\mu] = -K_\mu, \qquad [M_{\mu\nu}, P_\lambda] = \delta_{\nu\lambda} P_\mu - \delta_{\mu\lambda} P_\nu, \tag{25}$$

we have, using $M_{\mu\mu} = 0$,

$$\begin{aligned}
\langle \mathcal{O}|K_\mu K^\mu P_\nu P^\nu|\mathcal{O}\rangle &= \langle \mathcal{O}|K^\mu(2\delta_{\mu\nu} D - 2M_{\mu\nu} + P_\nu K_\mu)P^\nu|\mathcal{O}\rangle \\
&= \langle \mathcal{O}|2(1+\Delta)K_\nu P^\nu - 2K^\mu(dP_\mu - \delta_{\mu\nu}P_\nu - P^\nu M_{\mu\nu}) + 4(\delta_{\mu\nu}D - M_{\mu\nu})^2|\mathcal{O}\rangle \\
&= \langle \mathcal{O}|\mathcal{O}\rangle(4\Delta d(1+\Delta) - 4d^2\Delta + 4d\Delta + 4d\Delta^2) \\
&= 4d\Delta\langle \mathcal{O}|\mathcal{O}\rangle(2\Delta + 2 - d).
\end{aligned} \tag{26}$$

In order for this to be positive, we need either $\Delta = 0$ (which is the case if $\mathcal{O} = \mathbf{1}$, or $\Delta \geq (d-2)/2$, proving the bound. Note that since the inner product is positive-definite, if $\mathcal{O}$ is a scalar saturating the bound, then we necessarily have $\partial^2 \mathcal{O} = 0$, so that any scalar saturating the bound necessarily obeys the free-particle wave equation.

May 29 — From coupled Ising CFTs to the $O(2)$ model in three dimensions via conformal perturbation theory

Today we are doing an elaboration on an exercise given by Zohar Komargodski to the attendees of the 2017 Bootstrap school. The goal is to show that two 3d Ising models flow to

the $O(2)$ fixed point when coupled through a deformation $\epsilon_1\epsilon_2$. First we'll derive some needed conformal perturbation theory results, and then we'll attack the problem both through an ϵ expansion and a direct perturbation away from the 3d Ising CFT.

Solution:

General conformal perturbation theory

Consider a CFT perturbed by a collection of operators with dimensionless couplings λ_i and scaling dimensions Δ_i:

$$\delta S = \sum_i a^{-d+\Delta_i} \lambda_i \int \mathcal{O}_i, \tag{27}$$

where a is the short-distance cutoff (from here on, summation over repeated indices is implied).

The deformation will generically take the theory away from the fixed point. To get the β functions, we will choose an observable in the theory and require that its differential with respect to the short distance cutoff vanish. There are many observables to choose from, but the one we will find most convenient is the overlap between the state $|0\rangle$ and $|\mathcal{O}_i\rangle$ in the presence of the perturbation to the action in a region R. Therefore we will need to compute[7]

$$\frac{d}{d\ln a}\langle \mathcal{O}_i | e^{-\lambda_j a^{-d+\Delta_j} \int_R \mathcal{O}_j} |0\rangle = 0. \tag{28}$$

To evaluate this, we expand the exponential to quadratic order in the couplings, and using $\langle \mathcal{O}_i| = \lim_{x\to\infty}\langle 0|x^{2\Delta_i}\mathcal{O}(x)$, we have

$$\frac{d}{d\ln a}\left\langle 1 - V_R \lambda_i a^{-y_i} + \frac{\lambda_j\lambda_k}{2} a^{-y_j-y_k} \lim_{z\to\infty} \int_R d^d y\, d^d x\, \mathcal{O}^i(z)\mathcal{O}^j(x)\mathcal{O}^k(y) \right\rangle = 0, \tag{29}$$

with $y_i \equiv d - \Delta_i$ and V_R the volume of R. Here we used the OPE to simplify the second term. If we take R to be bounded we can do the OPE between $\mathcal{O}^j$ and $\mathcal{O}^k$, and then the OPE between the resulting operator and $\mathcal{O}^i$. Since we will be doing the OPE between an operator at ∞ and one in R, it doesn't really matter where exactly in R the latter operator is located. Therefore we can always take the OPE to be taken with x located at the center of R (we'll take R to be a ball). The integral over x then produces a factor of V_R. We then have remaining an integral $S^{d-1}\int dr\, r^{\Delta_i-\Delta_j-\Delta_k+d-1}$. We will absorb the S^{d-1} factor by re-scaling all the coupling constants, and so we then have, letting $d_t = \frac{d}{d\ln a}$ be the differential for RG time,

$$y_i a^{-y_i}\lambda_i - \beta_i a^{-y_i} + a^{-(y_j+y_k)}\left(\beta_j\lambda_k - \frac{\lambda_j\lambda_k}{2}(y_j+y_k-d_t)\right) C^i_{jk} \int dr\, \frac{1}{r^{-\Delta_i+\Delta_j+\Delta_k-d+1}} = 0 \tag{30}$$

[7]Other similar choices of observables, like $\langle e^{-\lambda_j a^{-d+\Delta_j}\int \mathcal{O}_j}\rangle$ or $\langle \mathcal{O}_i | e^{-\lambda_j a^{-d+\Delta_j}\int \mathcal{O}_j} | \mathcal{O}_k\rangle$, do not give us, upon differentiation, formulas which are as nice. This is because they have β functions mixed up with annoying integrals. The merits of the observable we chose to study lie mainly in the fact that the linear term in the coupling λ_i is nonzero—after differentiating this gives us a nice factor of β_i all by itself, and this makes the resulting manipulations easier.

Note that to obtain this equation, we had to assume that the theory the expectation value was being taken with respect to was a CFT—otherwise the d_t pick up extra terms from the change in the action being used to construct the expectation value.

To satisfy this equation to first order in the couplings, we need $\beta_i = y_i\lambda_i + O(\lambda^2)$. In the term $\beta_j\lambda_k$, only the linear part of β_j contributes at quadratic order in the couplings, and this term cancels the $-\lambda_j\lambda_k(y_j+y_k)/2$ piece. Therefore, writing the quadratic piece of β_i as $\beta^{(2)}$, we have

$$\beta_i^{(2)} a^{-y_i} = C_{jk}^i a^{-y_j-y_k}\frac{\lambda_j\lambda_k}{2} d_t \int dr \frac{1}{r^{-\Delta_i+\Delta_j+\Delta_k-d+1}}. \tag{31}$$

The differential of the integral gives $-d_t a^{\Delta_i-\Delta_j-\Delta_k+d}/(\Delta_i-\Delta_j-\Delta_k+d) = -a^{-y_i+y_j+y_k}$, and so

$$\beta_i^{(2)} = -\frac{\lambda_j\lambda_k}{2}C_{jk}^i \implies \beta_i = y_i\lambda_i - \frac{\lambda_j\lambda_k}{2}C_{jk}^i. \tag{32}$$

This gives us the β functions for the couplings of the perturbations added to the action to deform the theory away from the fixed point. To see the extent to which other operators (viz. those corresponding to the scaling variables in the CFT) are modified by the perturbation, we need to make a further deformation and include these operators in the exponential. For example, we will want to compute the extent to which the perturbation modifies the scaling dimension of an operator $\mathcal{O}$ in the CFT. This is done by adding $\delta\lambda_{\mathcal{O}}\int\mathcal{O}$ to the exponential in the deformation, and then expanding the exponential to linear order in $\delta\lambda_{\mathcal{O}} = \lambda_{\mathcal{O}} - \lambda_{\mathcal{O}*}$, where $\lambda_{\mathcal{O}*}$ is the value of the coupling in the CFT. This is because in order to determine the scaling dimension of $\mathcal{O}$ in the CFT, we need to perturb away from the CFT by deforming by $\mathcal{O}$, and then examining how quickly the theory flows back to the fixed point. To the extent that $\mathcal{O}$ remains a scaling variable in the presence of the $\lambda_i\mathcal{O}_i$ deformation, we then find that $\beta_{\mathcal{O}}$ is made nonzero by the term

$$\beta_{\mathcal{O}} = y_{\mathcal{O}}^{(0)}\delta\lambda_{\mathcal{O}} - (\delta\lambda_{\mathcal{O}})\lambda_j C_{\mathcal{O}j}^{\mathcal{O}}, \tag{33}$$

where $d - y_{\mathcal{O}}^{(0)}$ is the scaling dimension of $\mathcal{O}$ in the un-deformed theory. Since $\delta\lambda_{\mathcal{O}}$ vanishes at the fixed point, this means that the scaling dimension of $\mathcal{O}$ is corrected by

$$\Delta_{\mathcal{O}} = \Delta_{\mathcal{O}}^{(0)} + \lambda_{j*}C_{\mathcal{O}j}^{\mathcal{O}}. \tag{34}$$

ϵ expansion from $d = 4$

Now we consider two Ising models, coupled though the product of their energy operators. The action in $d = 4-\epsilon$ dimensions is

$$S = \int\left(\frac{1}{2}\partial\boldsymbol{\phi}\cdot\partial\boldsymbol{\phi} + ta^{-2}\boldsymbol{\phi}\cdot\boldsymbol{\phi} + ga^{-\epsilon}(\phi_1^4+\phi_2^4) + \eta a^{-\epsilon}\phi_1^2\phi_2^2\right). \tag{35}$$

We will evaluate the β functions around the Guassian fixed point, and use them to predict where the theory flows. The starting point of the flow is hence just free field theory, and so getting the OPE coefficients is but a simple matter of combinatorics. They are (normalizing the fields so that $\langle\phi_i(r)\phi_j(0)\rangle = \frac{1}{x^{d-2}}$)

$$C_{\eta\eta}^{\eta} = \binom{2}{1}^4 = 16, \quad C_{\eta g}^{\eta} = 2\cdot 2!\binom{4}{2} = 24, \quad C_{gg}^{g} = 2!\binom{4}{2}^2 = 72, \quad C_{\eta\eta}^{g} = 2. \tag{36}$$

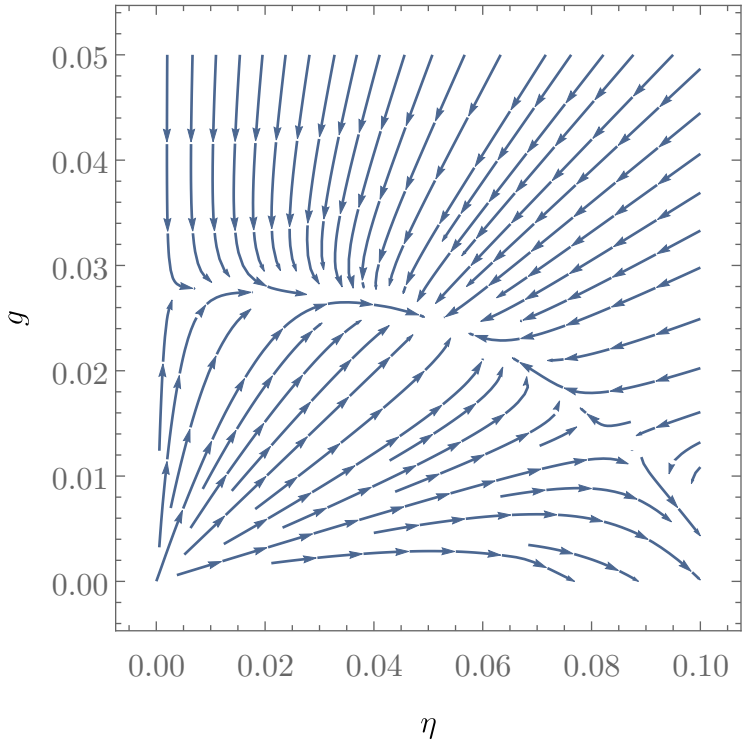
The ones involving t are

$$C^t_{tt} = \binom{2}{1}^2 = 4, \quad C^t_{gg} = \binom{4}{3}^2 3! = 96, \quad C^t_{gt} = \binom{4}{2} 2! = 12, \quad C^t_{\eta\eta} = \binom{2}{1}^2 2! = 8, \quad C^t_{\eta t} = 2$$
$$C^g_{tt} = 1, \quad C^\eta_{tt} = 2, \quad C^g_{tg} = \binom{4}{1} 2! = 8, \quad C^\eta_{\eta t} = 2 \cdot 2^2 = 8. \tag{37}$$

First let us ignore t—we tune it to zero in the UV, and then anticipate that the error in our fixed point value $t_* = 0$ will only be of order ϵ^2 (since the non-zero β function for t at $t = 0$ comes only from η^2, g^2 terms). Then plugging the relevant OPE coefficients in to our formulae for the β functions (all of t, g, η are being treated as perturbations), we obtain

$$\beta_\eta = \epsilon\eta - 8\eta^2 - 24\eta g, \qquad \beta_g = \epsilon g - \eta^2 - 36 g^2. \tag{38}$$

The flow looks like



(39)

There are four fixed points, only one of which is IR stable with respect to both η and g. We find that at this point,

$$\eta_* = \frac{\epsilon}{20}, \qquad g_* = \frac{\epsilon}{40}, \qquad y_\eta = -\epsilon, \qquad y_g = -\epsilon/5, \tag{40}$$

while the fixed point value of the quadratic term vanishes: $t_* = 0$. Therefore at the fixed point we may write the action, to the extent that doing so is meaningful, as

$$S = \int \left(\frac{1}{2} \partial\boldsymbol{\phi} \cdot \partial\boldsymbol{\phi} + \frac{\epsilon}{40} a^{-\epsilon} (\boldsymbol{\phi} \cdot \boldsymbol{\phi})^2 \right). \tag{41}$$

Note that the fixed-point values of η, g are such that the resulting action is $O(2)$-symmetric.

Now if we go back and include t, we have

$$\beta_t = 2t - 2t^2 - 2t\eta - 12tg - 48g^2 - 4\eta^2, \tag{42}$$

and

$$\beta_\eta = \beta_\eta(t=0) - t^2 - 8\eta t, \qquad \beta_g = \beta_g(t=0) - 8tg - \frac{1}{2}t^2. \tag{43}$$

When $\epsilon \ll 1$ we have $t_* \sim O(\epsilon^2)$, which can be dropped. We then get that at the fixed point, (this isn't quite correct since the presence of g^2, η^2 terms mean that t is not by itself a scaling variable at the fixed point, even to order ϵ)

$$y_t = 2 - 2\epsilon(1/20 + 6/40) + O(\epsilon^2) = 2(1 - \epsilon/5) + O(\epsilon^2). \tag{44}$$

When $\epsilon \sim 1$ the solutions for the fixed points get very complicated, and saying anything concrete gets more difficult. However, we note that setting $\epsilon = 1$ in the above equation (and yet still dropping the $O(\epsilon^2)$ terms, lol) gives a scaling dimension of $\Delta_t = 7/5 = 1.4$ for the energy operator in the $O(2)$ model. In fact the actual scaling dimension is $\Delta_t \approx 1.51$, so this laughably crude approximation actually gets the right scaling dimension within 7% error.

Perturbation theory direction from the Ising CFT

Now we'll take a more direct approach. Instead of starting from the Gaussian fixed point in 4 dimensions and flowing along the WF trajectory towards the putative $O(2)$ fixed point, we will start with two coupled Ising models in three dimensions, and flow along the trajectory generated by the inter-model coupling. Since the inter-model coupling term is only very slightly relevant, the flow will likely be a lot shorter than the flow from the Gaussian fixed point in 4 dimensions, and so will hopefully provide us with a better prediction for the $O(2)$ CFT data.

The Ising models will be coupled together through their energy operators:

$$S = S_{I,1} + S_{I,2} + \eta \int \epsilon_1 \epsilon_2. \tag{45}$$

First let us recall some results for the 3d Ising model, namely [**?**]

$$\Delta_\epsilon \approx 1.41, \qquad \Delta_{\epsilon'} \approx 3.82, \qquad C^\epsilon_{\epsilon\epsilon} \approx 1.53, \qquad C^{\epsilon'}_{\epsilon\epsilon} \approx 1.54, \tag{46}$$

where ϵ is the energy operator (schematically, ϕ^2) and ϵ' is the next lightest $\mathbb{Z}_2$-even scalar in the spectrum. Note that an operator like ϵ^2, with scaling dimension the relevant ≈ 2.8, doesn't appear in the spectrum. Also note that $[\epsilon_1\epsilon_2] \approx 2.82 < 3$, and so the deformation that couples the two Ising models at the decoupled fixed point $\mathcal{I}$ is slightly relevant: $y^{\mathcal{I}}_\eta \approx 0.18$. The beta function for the deformation is therefore

$$\beta_\eta = y^{\mathcal{I}}_\eta \eta - \frac{\eta^2}{2} C^\eta_{\eta\eta}, \qquad C^\eta_{\eta\eta} = (C^\epsilon_{\epsilon\epsilon})^2. \tag{47}$$

Therefore the IR fixed point $\mathcal{F}$ in this approximation is at

$$\eta_* = 2y^{\mathcal{I}}_\eta (C^\epsilon_{\epsilon\epsilon})^{-2} \approx 0.15, \tag{48}$$

and the dimension of η at the IR fixed point $\mathcal{F}$ is

$$\Delta_\eta^{\mathcal{F}} = d + y_\eta^{\mathcal{I}} \approx 3.18 \implies y_\eta^{\mathcal{F}} = -0.18, \tag{49}$$

Going back and calculating what we would expect for $y_\eta^{\mathcal{I}}$ at the fixed point with the ϵ expansion,[8] we obtain $y_\eta^{\mathcal{F}} = -2/5$, which is in pretty good agreement with the (presumably more accurate) result above.

We can also calculate the extent to which the coupling modifies the scaling dimension of the toatl energy operator $\boldsymbol{\epsilon}$, pretending that $\boldsymbol{\epsilon} \equiv \epsilon_1 + \epsilon_2$ is still an exact scaling variable at the IR fixed point. Using the results obtained in the first section, we see that the scaling dimension at the IR fixed point is approximately

$$\Delta_{\boldsymbol{\epsilon}}^{\mathcal{F}} = \Delta_\epsilon^{\mathcal{I}} + \eta_* C_{\eta\boldsymbol{\epsilon}}^{\boldsymbol{\epsilon}}. \tag{50}$$

The scaling dimension of the energy operator thus increases along the flow. Since $C_{\boldsymbol{\epsilon}\eta}^{\boldsymbol{\epsilon}} = C_{\epsilon\epsilon}^{\epsilon}$, we have

$$\Delta_{\boldsymbol{\epsilon}}^{\mathcal{F}} = \Delta_\epsilon^{\mathcal{I}} + 2\frac{3 - \Delta_\eta^{\mathcal{I}}}{C_{\epsilon\epsilon}^{\epsilon}} \approx 1.63. \tag{51}$$

The actual scaling dimension of $\boldsymbol{\epsilon}$ in the $O(2)$ is in fact $\Delta_{\boldsymbol{\epsilon}}^{O(2)} \approx 1.51$, which eh, isn't too bad for a leading-order approximation. Unfortunately it does about the same as the more crude ϵ expansion estimate, although it over-estimates $\Delta_{\boldsymbol{\epsilon}}$ instead of under-estimating, like the ϵ expansion does (in fact, the ϵ expansion predicts $\Delta_{\boldsymbol{\epsilon}} < \Delta_\epsilon$ (by a tiny bit), which we know can't be right).

May 31 — Supersymmetric localization and the geometry of phase space path integrals

Today we're doing a computation that illustrates the basic features of supersymmetric localization in a very simple geometric way. This is a slight elaboration on a problem that was assigned in David Skinner's SUSY course; I found the problem statement on his webpage.

Consider a symplectic manifold (M, ω) of dimension $2m$ equipped with a Hamiltonian map $H : M \to \mathbb{R}$ which generates a $U(1)$ action on M. Consider the zero-dimensional action (the factor of $1/2$ isn't in the problem statement I found online, but I think it's necessary)

$$S = \alpha\left(H(x) + \frac{1}{2}\omega_{ab}\psi^a\psi^b\right), \tag{52}$$

where the ψ^a are a set of $2m$ Grassmann variables. Consider the supersymmetry generator

$$Q = \psi^a\frac{\partial}{\partial x^a} + V^a(x)\frac{\partial}{\partial \psi^a}, \tag{53}$$

[8] Again, because of the $g\eta$ term in β_η, η is technically speaking not a scaling variable.

where V is the Hamiltonian vector field associated to the action of H. Use localization techniques to exactly compute the partition function

$$Z = \frac{1}{(2\pi)^m} \int d^{2m}x \, d^{2m}\psi \, e^{-S(x,\psi)}. \tag{54}$$

Solution:

First let's try to understand the SUSY operator Q better. We can get a better understanding of Q by computing its square:

$$Q^2 = V^a \frac{\partial}{\partial x^a} + \psi^a \frac{\partial V^b}{\partial x^a} \frac{\partial}{\partial \psi^b}. \tag{55}$$

In what follows, it will be helpful to think geometrically. This is done by realizing that the ψ^a variables are really just a way of writing /basis covectors dx^a. In this notation, a 1-form V is written as $V_a\psi^a$, which has all the properties of a 1-form written as $V_a dx^a$ (the weird way that fermion measures transform allows the ψ^a to behave as basis covectors). Thinking about this, we realize that $\psi^a \frac{\partial}{\partial x^a}$ is really just $d = dx^a \wedge \partial_a$, and that $V^a(x)\frac{\partial}{\partial \psi^a}$ is really just i_V, which operates by contraction: $i_V(B_a dx^a) = V^a B_a$, $i_V(C_{ab} dx^a \wedge dx^b) = V^a C_{ab} dx^b$, and so on. Therefore, in geometric language, the SUSY operator is the equivariant derivative associated with the vector field V, viz.

$$Q = d_V = d + i_V. \tag{56}$$

Now let us identify the square of Q. In differential geometry language, converting ψ^as to dx^as, we claim that

$$Q^2 = \mathcal{L}_V = i_V d + d i_V, \tag{57}$$

which is the Lie derivative along V. To show this, we can e.g. consider how Q^2 acts on a 1-form $\rho = \rho_a \psi^a = \rho_a dx^a$:

$$Q^2 \rho = \left(\frac{\partial V^a}{\partial x^b} \rho_a + V^a \frac{\partial \rho_b}{\partial x^a} \right) \psi^b. \tag{58}$$

Let us then compute

$$\mathcal{L}_V \rho = i_V(d\rho) + d(i_V \rho). \tag{59}$$

The first term is

$$i_V(d\rho) = i_V \partial_a \rho_b \psi^a \psi^b = V^a \left(\frac{\partial \rho_b}{\partial x^a} - \frac{\partial \rho_a}{\partial x^b} \right) \psi^b = V^a \partial_{[a}\rho_{b]} dx^b. \tag{60}$$

The second term is

$$d(i_V \rho) = \frac{\partial V^a}{\partial x^b} \rho_a \psi^b + V^a \frac{\partial \rho_a}{\partial x^b} \psi^b. \tag{61}$$

Adding these two together, we get

$$\mathcal{L}_V \rho = \left(\frac{\partial V^a}{\partial x^b} \rho_a + V^a \frac{\partial \rho_b}{\partial x^a} \right) \psi^b = Q^2 \rho, \tag{62}$$

and so Q^2 indeed acts as the Lie derivative.

Let's now show that the action is invariant under SUSY transformations. The first term transforms simply as

$$QH = \psi^a \frac{\partial H}{\partial x^a}, \tag{63}$$

which in differential geometry would read $QH = d_V H = dH$. The second term is

$$Q\left(\frac{1}{2}\omega_{ab}\psi^a\psi^b\right) = \frac{1}{2}\left(\psi^c \frac{\partial \omega_{ab}}{\partial x^c}\psi^a\psi^b + \omega_{ab}V^a\psi^b - \omega_{ab}\psi^a V^b\right) = \omega_{ab}V^a\psi^b, \tag{64}$$

where we have used the antisymmetry of ω and the fact that ω is closed as a 2-form, so that the antisymmetrization enacted by the fermions in the first term makes the first term vanish. In terms of differential geometry language, the above computation reads

$$Q\omega = d_V\omega = d\omega + i_V\omega = V^a\omega_{ab}dx^b. \tag{65}$$

Putting these two transformations together, we see that the action is indeed invariant under SUSY, since

$$QS = \psi^a \frac{\partial H}{\partial x^a} + \omega_{ab}V^a\psi^b = \psi^a\omega_{ab}V^b + \omega_{ab}V^a\psi^b = 0, \tag{66}$$

where we've used the definition of the Hamiltonian vector field V, namely that V is determined by dual to $\frac{\partial}{\partial x^a}H$ via the symplectic form, which sets up the isomorphism $\omega : TM \cong T^*M$:

$$V^a = \omega^{ab}\frac{\partial H}{\partial x^b}. \tag{67}$$

Just for fun, let's recall why this is true: a Hamiltonian vector field V is a vector field in phase space along which the symplectic form is preserved, which means that

$$0 = \mathcal{L}_V\omega = d(i_V\omega), \tag{68}$$

since $d\omega = 0$. Therefore locally $i_V\omega$ is exact, and we can write it as $i_V\omega = dH$, where H is of course the Hamiltonian.

In order to run the localization procedure, we will add a Q-exact term to the action, and then show that the path integral is unchanged by its presence. The natural Q-exact term to add is the equivariant derivative $d_V V = QV$. More formally, let g be a $U(1)$-invariant metric on M.[9] Define $\Psi = g(V, \psi) = \psi^a V_a$, and consider adding $Q\Psi$ to the action, by defining[10]

$$S_\lambda(x, \psi) \equiv S(x, \psi) + \lambda Q\Psi, \tag{69}$$

where

$$Q\Psi = V^a V_a + \psi^a\psi^b\partial_a V_b. \tag{70}$$

[9] If we are given a metric that's not $U(1)$-invariant, we can get an invariant one just by averaging over the $U(1)$ action, so this is done wolog.

[10] I'm getting sick of the fractions, so from now on ∂_a means differentiation wrt x^a (fermionic derivatives will be expressed more verbosely).

Now we will show that $Q^2\Psi = 0$, which will mean that $QS_\lambda = 0$. Based on the geometric interpretation of Q that we just gave, the differential-geometric way to prove this is is to show that

$$Q^2\Psi = \mathcal{L}_V V = (i_V d + d i_V)V \tag{71}$$

vanishes. Indeed, the two terms on the RHS are

$$i_V dV = i_V(\partial_a V^b dx^a \wedge dx^b) = V^a \partial_{[a} V_{b]} dx^b = (V^a \partial_a V_b - V^a \partial_b V_a) dx^b, \qquad d i_V V = 2V^a \partial_b V_a dx^b, \tag{72}$$

and so

$$\mathcal{L}_V V = V^a(\partial_a V_b + \partial_b V_a)dx^b = 0, \tag{73}$$

since V is Killing field.

Now let us consider the variation of the partition function Z_λ with respect to λ:

$$(2\pi)^m \partial_\lambda Z_\lambda = \int d^{2m}x\, d^{2m}\psi\, \partial_\lambda S_\lambda e^{-S_\lambda} = \int d^{2m}x\, d^{2m}\psi\, Q(\Psi) e^{-S_\lambda} = \int d^{2m}x\, d^{2m}\psi\, Q\left(\Psi e^{-S_\lambda}\right), \tag{74}$$

since $QS_\lambda = 0$. However, both terms that are created when Q acts on something are total derivatives—either a total ψ derivative (in which case we get zero since $\int d\psi = 0$) or a total x^a derivative (in which case we get zero if $H(x) \to \infty$ at the limits of x integration, which we assume). Therefore $\partial_\lambda Z_\lambda = 0$, and so Z_λ is actually independent of λ. This means that we can send $\lambda \to \infty$ without changing the partition function. Now the exponential of the $\lambda\psi^a\psi^b\partial_a V_b$ term, being fermionic, has a power series expansion that truncates after a finite number of terms and ensures that its contribution to the partition function occurs as λ^m. The exponential of $-\lambda V^a V_a$ will thus dominate over the λ^m term and send the integrand to zero, unless x is such that $V^a V_a(x) = 0$. Let x_* denote a point where $V^a(x_*) = 0$. Since the symplectic form is non-degenerate, this must mean that at x_* we have $(\partial_a H)(x_*) = 0$, and so the points that the integral localizes around are the critical points of the Hamiltonian.

Another way to see that the integral localizes, which doesn't rely on adding the extra term to the action, is to realize that anything which is Q-closed is also Q exact everywhere except for the critical points where $V(x_*) = 0$. Since $Qe^{-S} = 0$ the integrand in Z is Q-closed, and so this means that the integrand is a total derivative everywhere except the critical points; by Stoke's theorem we then see that Z receives contributions only from the critical points.

To prove the above claim about Q-exactness, we proceed as follows. Consider the field (best thought of as an inhomogeneous differential form in $\Omega^\bullet(M)$)

$$\Gamma = V \wedge (QV)^{-1}, \tag{75}$$

which is defined on $M \setminus X_*$, where X_* is the set of critical points. Here the inverse of QV is defined as (note to self: from here until when I go back to discussing the $\lambda Q\Psi$ tactic, the operator Q is $d - i_V$ instead of $d + i_V$ as above—bleh, at least it doesn't affect the conclusions. Will cleanup later)

$$(QV)^{-1} = (dV - V^2)^{-1} = -\frac{1}{V^2}(1 - dV/V^2)^{-1} = -\frac{1}{V^2}\sum_{i=1}^{m} V^{-2k}(dV)^{\wedge k}, \tag{76}$$

where again the notation is $V = V_a\psi^a$, $dV = \partial_a V_b \psi^a\psi^b/2$. This is the proper way to take an inverse of an inhomogeneous differential form for the same reason that $\sum_j (-x)^j = (1+x)^{-1}$. In our case the sum is finite and so we get

$$(1 - dV/V^2)\sum_{j=1}^{m}(dV/V^2)^{\wedge j} = 1 - (dV/V^2)^{\wedge(m+1)}, \tag{77}$$

but the last term vanishes since its dimension is too large. Note that the inverse is only well-defined away from the critical points.

A more pedestrian approach, which convinces us that this is the right way to take the inverse, is as follows. We want to find Γ such that $Q\Gamma = 1$. Since Q is odd, we can parametrize Γ as a series of odd terms:

$$\Gamma = -\frac{A_a\psi^a}{V^2} - \frac{B_{abc}\psi^a\psi^b\psi^c}{V^2} - \dots, \tag{78}$$

where the minus signs and $1/V^2$s are just for convenience. Now the first term, when acted on by Q, is

$$Q\Gamma \ni \frac{V^a A_a}{V^2} - \frac{\partial_a A_b \psi^a\psi^b}{V^2} + \frac{2}{V^4}V^c\partial_a V_c A_b \psi^a\psi^b. \tag{79}$$

Since the first term is the only one in $Q\Gamma$ with no ψs, we need $A_a = V_a$, so that

$$Q\Gamma \ni 1 - \frac{\partial_a V_b \psi^a\psi^b}{V^2} + \frac{2}{V^4}V^c\partial_a V_c V_b \psi^a\psi^b. \tag{80}$$

The second term in Γ, when acted on by Q, contributes 2- and 4-fermion terms. If $Q\Gamma = 1$, then we need the 2-fermion terms to precisely the 2-fermion terms in the above equation. The 2-fermion terms produced by the B term in $Q\Gamma$ are

$$Q\Gamma \ni \frac{V^a}{V^2}\left(-B_{abc} + B_{bac} - B_{bca}\right)\psi^b\psi^c. \tag{81}$$

Suppose we choose $B_{abc} = V^{-2}V_a\partial_b V_c$, which is what the above formula for the inverse tells us to do. Then the above terms become

$$Q\Gamma \ni \frac{1}{V^4}\left(V^2\partial_a V_b - V_a V^c\partial_c V_b + V_a V^c\partial_b V_c\right)\psi^a\psi^b. \tag{82}$$

Therefore we have, re-naming some dummy indices and letting $\dots$ denote terms with 4 or more fermions,

$$Q\Gamma = 1 - \frac{1}{V^4}\left(V_a V^c\partial_b V_c + V_a V^c\partial_c V_b\right) + \dots = 1 - \frac{V^a V_c}{V^4}\partial_{(b}V_{c)} = 1 + \dots \tag{83}$$

since V is Killing.

One can then check that choosing the next term to be $-V^{-2}C_{abcde}\psi^a\cdots\psi^e$ with $C = (dV\wedge dV)/V^4$ cancels the 4-fermion terms in the QB_{bcd} term, and that the 6-fermion terms in the QC_{abcde} term are canceled by the next order term, and so on. This cancellation occurs up until we reach products of m fermions, at which point all further terms die by antisymmetry.

Anyway, using Γ we can note that since $Q\Gamma = 1$,

$$QS = 0 \implies Q(\Gamma S) = S, \tag{84}$$

and so S is Q-exact. Therefore all Q-closed forms are exact away from the critical points, proving our claim about localization.

Anyway, going back to the $\lambda Q\Psi$ approach which is easier to implement in practice, we expand about the critical point and then define y by $y = x\sqrt{\lambda}$ and $\widetilde{\psi}$ by $\widetilde{\psi} = \psi\sqrt{\lambda}$,[11] yielding

$$\begin{aligned} Z = Z_{\lambda\to\infty} = \frac{1}{(2\pi)^m} \sum_{x_*:V^a(x_*)=0} \int d^{2m}y\, d^{2m}\psi\, \exp\Big[\alpha\Big(- H(x_*) \\ - \frac{1}{2}(\partial_a\partial_b(V_cV^c))(x_*)y^ay^b - (\partial_{[a}V_{b]})(x_*)\widetilde{\psi}^a\widetilde{\psi}^b + \dots\Big), \end{aligned} \tag{85}$$

where $\dots$ are terms that vanish when $\lambda \to \infty$ and hence can be dropped. We can now do the integrals no problem:

$$Z = \sum_{x_*:V(x_*)=0} e^{-\alpha H(x_*)}\alpha^m \frac{\mathrm{Pf}[(\partial_{[a}V_{b]})(y_*)]}{\sqrt{\det[(\partial_a\partial_bV_cV^c)(y_*)]}}. \tag{86}$$

Now around each critical point, we can choose coordinates where V has the canonical form of a Hamiltonian vector field, viz.

$$V(x_*) = \sum_{i=1}^m n_i(x_*)\left(p^i\frac{\partial}{\partial q^i} - q^i\frac{\partial}{\partial p^i}\right), \tag{87}$$

where we have split up the coordinates near x_* into a set of m "coordinates" q_i and m "momenta" p_i. Here the $n_i \in \mathbb{Z}$ since H generates a $U(1)$ symmetry. With this choice of coordinates, about each critical point we have

$$\mathrm{Pf}[(\partial_{[a}V_{b]})(x_*)/2] = \prod_{i=1}^m n_i(x_*), \qquad \det[(\partial_a\partial_bV_cV^c)(x_*)/2] = \det[(\partial_aV_c\partial_bV^c)(x_*)] = \prod_{i=1}^m n_i^4(x_*), \tag{88}$$

and so

$$Z = \sum_{y_*:V(y_*)=0} e^{-\alpha H(y_*)}\frac{\alpha^{-m}}{\prod_{i=1}^m n_i(y_*)}. \tag{89}$$

Finally, we can massage this by noting that the fermion path integral in the $\lambda = 0$ action can also be done exactly, giving

$$\int d^{2m}\psi\, e^{-\frac{1}{2}\omega_{ab}\psi^a\psi^b} = \frac{1}{m!}\omega^{\wedge m}, \tag{90}$$

which is the usual phase-space measure (on the LHS $\omega = \omega_{ab}dx^a \wedge dx^b$). Therefore the localization formula is

$$\frac{1}{m!}\int \omega^{\wedge m}\, e^{-\alpha H(x)} = \sum_{x_*:V(x_*)=0} e^{-\alpha H(x_*)}\frac{(2\pi/\alpha)^m}{\prod_{i=1}^m n_i(x_*)}. \tag{91}$$

[11] The integration measure is invariant since the fermionic and bosonic measures transform oppositely.

June 1 — Fun with Majorana quantum mechanics, path integrals, and traces of γ matrix exponentials

Today we're doing another problem from David Skinner's SUSY class. This one is pretty easy and shares a few features in common with a previous diary entry on Majorana path integrals, so we will be rather succinct.

Consider the QM action

$$S[\psi] = \frac{1}{2}\int_{S^1} d\tau \left(\psi_a \partial_\tau \psi^b + \omega_{ab}\psi^a\psi^b\right), \tag{92}$$

where ω is antisymmetric and $a = 1, \ldots, 4$. Using the path integral obtained from this action, find expressions for

$$\mathrm{Tr}[\bar{\gamma} e^{-\omega_{ab}\gamma^a\gamma^b/4}], \qquad \mathrm{Tr}[e^{-\omega_{ab}\gamma^a\gamma^b/4}] \tag{93}$$

in terms of sinh and cosh expressions. Here $\{\gamma^a, \gamma^b\} = 2\delta^{ab}$ are the 4d Gamma matrices.

Solution:

First, some preliminaries. Consider an operator $\mathcal{O}$ acting on a single two-dimensional Fock space. Then we claim that

$$\mathrm{Tr}[\mathcal{O}] = \int d\bar{\eta}\, d\eta\, e^{-\bar{\eta}\eta}\langle -\bar{\eta}|\mathcal{O}|\eta\rangle, \tag{94}$$

where the coherent states are, in our conventions,

$$|\eta\rangle = e^{-\eta\psi^\dagger}|0\rangle, \qquad \langle\bar{\eta}| = \langle 0|e^{\bar{\eta}\psi}, \tag{95}$$

with $\psi^\dagger|0\rangle = |1\rangle$ (the ψs are operators; the ηs are Grassmann numbers). With these conventions, $\langle\bar{\eta}|\eta\rangle = e^{\bar{\eta}\eta} = 1 + \bar{\eta}\eta$. One can then check that the minus sign in the left bra in (94) is needed to get the trace, and that if it's replaced with a $\langle\bar{\eta}|$, one instead gets the supertrace.

When we Trotterize the above expression for the trace with $\mathcal{O} = e^{-\beta H}$ (remembering the $e^{-\bar{\eta}_t\eta_t}$ factors!), we see that we get the usual $\bar{\eta}\partial_t\eta - H$ action, provided that we can identify $\bar{\eta}_N(\eta_1 + \eta_N)/\delta\tau$ with a time derivative (where there are N steps in the Trotterization). Therefore after Trotterizing, the trace computes the path integral with antiperiodic boundary conditions. Similarly, the supertrace $\mathrm{Tr}[(-1)^F\mathcal{O}]$ computes the trace with periodic boundary conditions.

To connect with the gamma matrices, one just realizes that four majoranas, when quantized, obey the same algebra as 4-dimensional gamma matrices. More precisely, because of normalization, the identification is $\psi^a = \gamma^a/\sqrt{2}$. Since $\bar{\gamma}$ anticommutes with all of them, we have in $(-1)^F = \bar{\gamma}$ in this representation. Therefore the partition functions with the two choices of boundary conditions are

$$Z_P = \mathrm{Tr}[\bar{\gamma} e^{-\omega_{ab}\gamma^a\gamma^b/4}], \qquad Z_A = \mathrm{Tr}[e^{-\omega_{ab}\gamma^a\gamma^b/4}]. \tag{96}$$

So, to compute these traces, we can alternately compute

$$Z_{A/P} = \det\left[\frac{1}{2}\frac{d}{d\tau} - \omega/2\right]^{1/2}, \tag{97}$$

with the eigenvalues for d_τ being determined from the boundary conditions.

First consider the case when the fermions are periodic around the S^1. Then we have

$$Z_P = \det\left[\prod_{n\in\mathbb{Z}}(i\pi n + \omega/2)\right]^{1/2}, \tag{98}$$

where now the determinant is taken only in the spinor tensor factor. Therefore

$$\mathrm{Tr}[\bar{\gamma}e^{-\omega_{ab}\gamma^a\gamma^b/4}] = \det\left[\frac{\omega}{2}\prod_{n\in\mathbb{N}}(\pi^2n^2 + \omega/2)\right]^{1/2} = N\det[\sinh(\omega/2)]^{1/2}, \tag{99}$$

where we've used

$$\sinh(x) = x\prod_{n\in\mathbb{N}}(1 + x^2/(\pi n)^2), \tag{100}$$

and where N is an ω-independent constant (determined e.g. through ζ regularization).

The antiperiodic spin structure case is basically the same: we get

$$\mathrm{Tr}[e^{-\omega_{ab}\gamma^a\gamma^b/4}] = \det\left[\prod_{n\in\mathbb{N}+1/2}(\pi^2n^2 + \omega/2)\right]^{1/2} = M\det[\cosh(\omega/2)]^{1/2}, \tag{101}$$

where M is another ω-independent constant. Here the relevant product formula is

$$\cosh(x) = \prod_{n\in\mathbb{N}}\left(1 + \frac{x^2}{(\pi n - \pi/2)^2}\right). \tag{102}$$

June 4 — Inversions, reflections, and SCTs

This is a short problem posed by Rychkov during his 2019 TASI lectures: show that a CFT has inversion symmetry iff it has reflection symmetry by relating the two symmetries with conjugation by an element of $SO(d+1,1)$.

Solution:

Since reflections (and inversions) aren't in the identity component of $O(d+1,1)$, CFTs don't necessarily have to have them as symmetries—the only symmetry a CFT must for sure have is $SO(d+1,1)$.

Our goal is to show that

$$R_\mu = O_\mu^{-1} I O_\mu, \qquad O_\mu \in SO(d+1,1), \tag{103}$$

where $I : x^\mu \mapsto x^\mu/x^2$ is the inversion and $R_\mu : x^\nu \mapsto (-1)^{\delta_{\mu,\nu}} x^\nu$ is the reflection of the μ coordinate. If this is true, then we will have proved that reflections and inversions are continuously connected—they are in the same connected component of $O(d+1,1)$—and so a CFT which has one as a symmetry necessarily has the other as a symmetry.

The inversion exchanges zero and infinity, and since R_μ leaves both zero and infinity invariant, any candidate O_μ will need to act nontrivially at infinity. The only generator of $SO(d+1,1)$ that acts at infinity is the SCTs K_ν, and so we will need to make use of the SCTs in constructing O_μ.

First let's try to understand what the hell the SCTs actually do. A finite SCT acts on the coordinates as

$$x^\mu \mapsto \frac{x^\mu - a^\mu x^2}{1 - 2a\cdot x + a^2 x^2}. \tag{104}$$

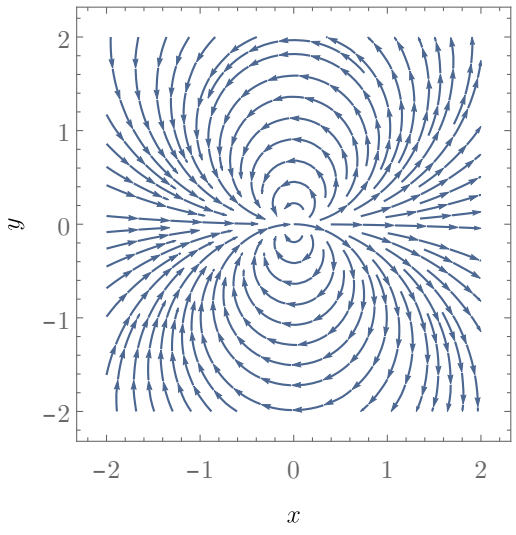
We can already see that this will move infinity, since it sends infinity to $-a^\mu/a^2$, and a^μ/a^2 to infinity. It also leaves zero invariant. In order to visualize better how it works, we can find the vector field associated with infinitesimal SCTs. Taking $|a|$ infinitesimal in the above equation, we find

$$x^\mu \mapsto 2(a\cdot x)x^\mu - a^\mu x^2 = a^\lambda (2x_\lambda x^\nu \partial_\nu - x^2 \partial_\lambda) x^\mu, \tag{105}$$

and so K_μ is associated with the vector field

$$K_\mu = 2x_\mu (x\cdot\partial) - x^2 \partial_\mu. \tag{106}$$

From this equation, we can see that K_μ acts by "swirling" around the origin. For example, in two dimensions, the operator K_x acts to move points along the following flow lines:



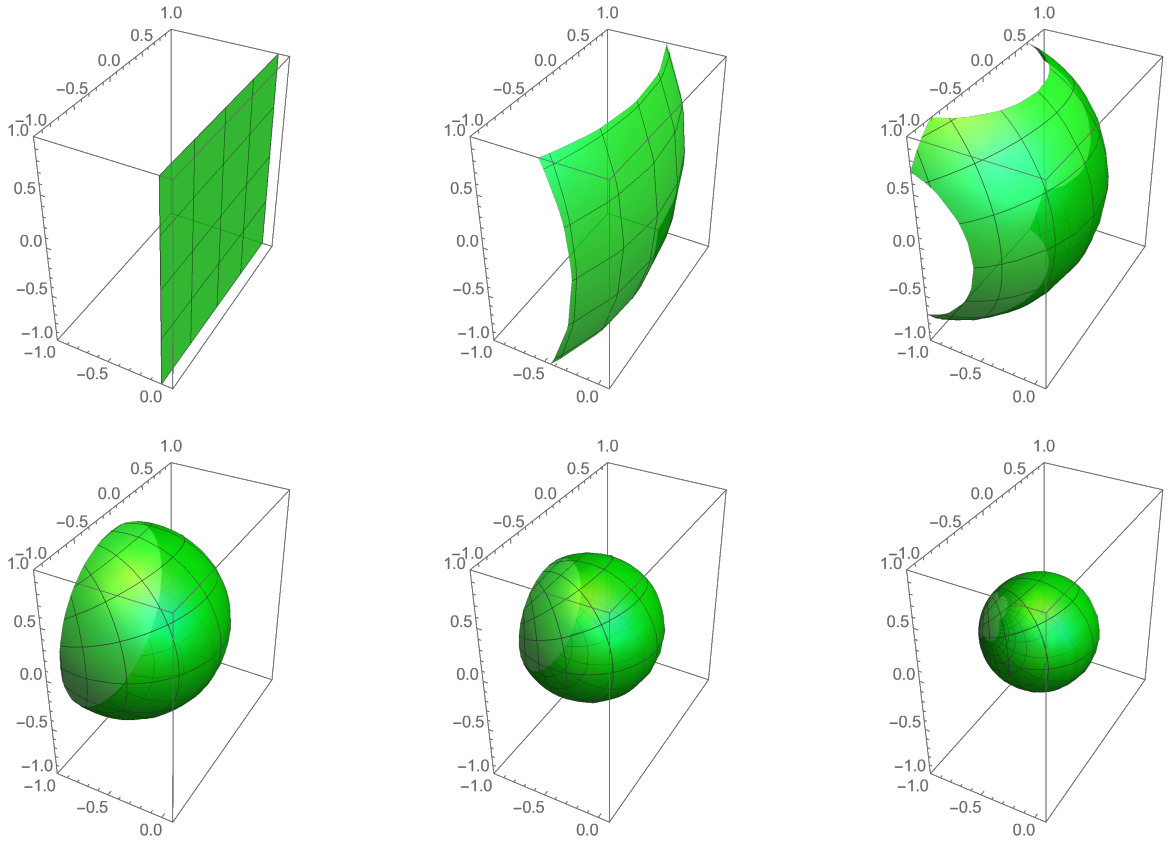
. (107)

I wish this plot had been included in e.g. the big yellow book; it was only after seeing this that I finally got some intuition for what K does.

Now, in order for $O_\mu^{-1} I O_\mu$ to be a reflection, O_μ must send the $x^\mu = 0$ hyperplane to the unit sphere, since the unit sphere is what I reflects about. The O_μ^{-1} will then map the sphere back to the $x^\mu = 0$ plane, and we will have accomplished a reflection about the plane. To see why the SCT can help do this, consider e.g. the image of the y axis in $\mathbb{R}^2$ under the SCT $K_x(1)$ (here our notation is that $K_\mu(t)$ is the SCT given by (104) with $a^\mu = t\hat{x}^\mu$):

$$K_\mu(1) : (0, y) \mapsto \frac{1}{1+y^2}(-y^2, y). \tag{108}$$

One then checks that the RHS defines a circle of radius 1 centered at $(-1/2, 0)$. This also happens in higher dimensions, with $K_\mu(1)$ mapping the $x^\mu = 0$ plane to the unit sphere centered on $x^\mu = -1/2$. This is perhaps best illustrated by some graphics. In the following figure, as me move along the grid, we show the image of the yz plane under $K_x(t)$ for t ranging from 0 to 1. By the time we're done, the plane is indeed compactified to a unit sphere centered at $(-1/2, 0, 0)$:



(109)

Anyway, summing up, we see that the reflection R_μ is smoothly connected to the reflection through the homotopy

$$R_\mu(t) = [T_{-\mu}(t/2) K_\mu(t)]^{-1} I T_{-\mu}(t/2) K_\mu(t), \tag{110}$$

where $t \in [0, 1]$ with $t = 1$ the reflection and $t = 0$ the inversion, and where $T_{-\mu}(t/2)$ is the translation by $t/2$ in the $-\hat{x}^\mu$ direction.

I think that this also provides us with a way to argue that any theory which is reflection-positive with respect to reflection about a plane is also reflection-positive in radial quantization (this works for primaries, but non-primary operators transform in more complicated ways under the mapping described above, though).

June 5 — Kahler pre-quantization practice

Today we're doing a long elaboration on a homework problem assigned by Greg Moore during his 2019 TASI lectures. We will give a prescription for doing pre-quantization of functions on phase space, and will then apply our machinery to the case where the phase space is S^2, thereby obtaining the Hilbert space for a quantum spinful particle.

Solution:

Generalities

First, some notation. Consider a phase space $\mathscr{P}$ with sympletic form ω. Functions[12] $f : \mathscr{P} \to \mathbb{C}$ are quantized by associating them to $\mathbb{C}$-linear operators $\mathcal{Q}[f]$ which act on $\mathcal{H}$. There are many ways of doing this, and so we need to fix some rules that determine what properties we would like the $\mathcal{Q}[f]$ to satisfy.

First, we will demand that commutation relations of the quantized functions are determined via the Poisson bracket as[13]

$$[\mathcal{Q}[f], \mathcal{Q}[g]] = i\hbar \mathcal{Q}[\{f, g\}], \tag{111}$$

where the Poisson bracket is as usual

$$\{f, g\} = \omega(f, g) = \omega^{ij}\partial_i f \partial_j g. \tag{112}$$

We could consider higher $\hbar$ corrections to (111) (deformation quantization), but for our purposes we will be able to find $\mathcal{Q}[f]$ such that (111) is satisfied exactly.

Now in this scheme, it is in general not possible to have the $\mathcal{Q}[f]$ operators form a linear representation; instead we require the weaker condition

$$||\mathcal{Q}[f]\mathcal{Q}[g] - \mathcal{Q}[f \cdot g]|| = O(\hbar^2). \tag{113}$$

Finally, we require that constants map as

$$\mathcal{Q}[a] = a\mathbf{1}, \tag{114}$$

and that complex conjugation of functions maps to Hermitian conjugation of operators:

$$\mathcal{Q}[f^*] = \mathcal{Q}[f]^\dagger. \tag{115}$$

[12] Really, sections of complex line bundles over $\mathscr{P}$.

[13] Note the sign on the RHS! Many references have $-i\hbar$ instead.

First, let's do a warmup, where we take $\mathscr{P} = \mathbb{R}^2$. Of course $\mathbb{R}^2 \cong T^*\mathbb{R}$ can be treated by sending $\mathcal{Q}[x^1] = x, \mathcal{Q}[x^2] = -i\hbar\partial_x$, but this quantization procedure is rather asymmetric in $\mathbb{R}^2$ and won't generalize to cases where the phase space is not a cotangent bundle, like $\mathscr{P} = S^2$. We will therefore proceed in complex coordinates, with symplectic form $\omega_{z\bar{z}} = -\omega_{z\bar{z}} = \frac{i}{2}$, so that $\omega = dx \wedge dy$. Therefore the Poisson bracket gives

$$\{z, \bar{z}\} = \omega^{z\bar{z}} = -2i, \tag{116}$$

which tells us that we ought to have

$$[\mathcal{Q}[z], \mathcal{Q}[\bar{z}]] = i\hbar\{z, \bar{z}\} = 2\hbar. \tag{117}$$

Consider then the assignment

$$\mathcal{Q}[z] = 2\hbar\bar{\partial} + z/2, \qquad \mathcal{Q}[\bar{z}] = -2\hbar\partial + \bar{z}/2. \tag{118}$$

This gives the correct commutator and satisfies $\mathcal{Q}[z]^\dagger = \mathcal{Q}[\bar{z}]$, and so we have successfully quantized the variables on $\mathscr{P}$.

For quantization on the flat plane, we could basically guess the form for the $\mathcal{Q}$ map—for more complicated examples, we will need to have a better system for figuring out what $\mathcal{Q}$ should be. In fact there are several options; quantization is not a unique procedure. The one I've found to be the most useful is to take

$$\begin{aligned} \mathcal{Q}[f] &= -i\hbar X_f^\mu \partial_\mu + \theta_\mu X_f^\mu + f \\ &= -i\hbar\nabla_{X_f} + \theta(X_f) + f \\ &= -i\hbar X_f^\mu D_\mu + f, \end{aligned} \tag{119}$$

where the covariant derivative is $D_\mu = \partial_\mu + i\theta_\mu/\hbar$ and X_f is the Hamiltonian vector field associated with f. Here our conventions are such that

$$X_f^\mu = \omega^{\mu\nu}\partial_\nu f \implies i_{X_f}\omega = -df. \tag{120}$$

Honestly this sign choice is sub-ideal, but this was only realized after finishing the calculations. The $+f$ term here is needed so that $\mathcal{Q}[1] = \mathbf{1}$, while the potential term $\theta(X_f)$ is needed for (111) to be satisfied, as we will see. Note that the sign on the potential is positive—the geometric quantization reference I was reading had the opposite sign, which led to a huge effort to track down where my calculation was going wrong, so use this sign when in doubt.

Note that with this choice of quantization we will never have $\mathcal{Q}[f \cdot g] = \mathcal{Q}[f]\mathcal{Q}[g]$ for f, g with $X_f, X_g \neq 0$, since the RHS is quadratic in derivatives while the LHS is linear. This is okay though, since the term quadratic in derivatives is always $O(\hbar^2)$.

The main reason for choosing this quantization is that it satisfies (111) on the nose (i.e. there are no higher $\hbar$ corrections). One drawback is that we can only ever produce operators that are linear in derivatives, which prevents us from getting formulae like e.g. $\mathcal{Q}[p^2] = -\hbar^2\nabla^2$. Sorting out this issue is a lot of work; I might come back to it at a later point.

Before proving why this quantization works, let's run a sanity check on the plane. We calculate

$$X_z = 2id\bar{z}, \qquad X_{\bar{z}} = -2idz, \qquad \theta = \frac{i}{4}(zd\bar{z} - \bar{z}dz), \tag{121}$$

so that

$$\theta(X_z) = -z/2, \qquad \theta(X_{\bar{z}}) = -\bar{z}/2. \tag{122}$$

Plugging these into (119), we recover the same quantization assignements as written above: for example, $\mathcal{Q}[z] = -i\hbar(2i)\bar{\partial} - z/2 + z = 2\hbar\bar{\partial} + z/2$ ✓.

Now we will prove that this choice of $\mathcal{Q}$ indeed satisfies (111) exactly. In the computation of $[\mathcal{Q}[f], \mathcal{Q}[g]]$, we have two groups of terms. The first comes from the commutator of the covariant derivatives, and gives

$$[X_f^\mu(-i\hbar\partial_\mu + \theta_\mu), X_g^\nu(-i\hbar\partial_\nu + \theta_\nu)] = -\hbar^2(\nabla_{X_f}X_g^\nu - \nabla_{X_g}X_f^\nu)\partial_\nu - i\hbar(\nabla_{X_f}\theta(X_g) - \nabla_{X_g}\theta(X_f)). \tag{123}$$

The second term gives, after taking the derivatives of $\theta(X_{f/g})$,

$$i\hbar(\nabla_{X_f}\theta(X_g) - \nabla_{X_g}\theta(X_f)) = i\hbar\omega(X_f, X_g) + i\hbar\theta(\nabla_{X_f}X_g - \nabla_{X_g}X_f), \tag{124}$$

since $\omega = d\theta$. Now as we will show in tomorrow's diary entry, for any two Hamiltonian vector fields X_f, X_g, we have

$$[X_f, X_g] = -X_{\{f,g\}}, \tag{125}$$

where the minus sign could be fixed if we went back and changed conventions to $i_{X_f}\omega = +df$. Therefore[14]

$$\begin{aligned}[X_f^\mu(-i\hbar\partial_\mu + \theta_\mu), X_g^\nu(-i\hbar\partial_\nu + \theta_\nu)] &= +\hbar^2\nabla_{X_{\{f,g\}}} + i\hbar\theta(X_{\{f,g\}}) - i\hbar\{f,g\} \\ &= i\hbar(-i\hbar\nabla_{X_{\{f,g\}}} + \theta(X_{\{f,g\}}) - \{f,g\}).\end{aligned} \tag{127}$$

This is almost right, except the curvature contribution $\{f,g\}$ has the wrong sign. This is fixed by the remainder of $[\mathcal{Q}[f], \mathcal{Q}[g]]$, which contains derivatives of f and g. This part is

$$[\mathcal{Q}[f], \mathcal{Q}[g]] \supset -i\hbar(\nabla_{X_f}g - \nabla_{X_g}f) = -i\hbar(X_f^\mu X_g^\nu\omega_{\mu\nu} - X_g^\mu X_f^\nu\omega_{\mu\nu}) = 2i\hbar\{f,g\}. \tag{128}$$

This flips the sign of the last term in (127), and so we get

$$[\mathcal{Q}[f], \mathcal{Q}[g]] = i\hbar(-i\hbar\nabla_{X_{\{f,g\}}} - \theta(X_{\{f,g\}}) + \{f,g\}) = i\hbar\mathcal{Q}[\{f,g\}], \tag{129}$$

as claimed.

[14]Here we need to pay attention to signs to determine that

$$\omega(X_f, X_g) = \omega_{\mu\nu}X_f^\mu X_g^\nu = X_f^\mu\partial_\mu g = \omega^{\mu\lambda}\partial_\lambda f\partial_\mu g = -\{f,g\}. \tag{126}$$

$\mathscr{P} = S^2$ and the quantum spin

Now let's specialize to the case where $\mathscr{P} = S^2$, with sympletic form $\omega \propto d\cos\theta \wedge d\phi$. To facilitate calculating Poisson brackets, we will find it useful to work in stereographic projection: looking back at the diary entry on the Hopf fibration, the coordinates in $\mathbb{R}^3$ are determined in terms of the complex coordinates as (here we are projecting from the north pole)

$$X = \frac{z+\bar{z}}{1+|z|^2}, \qquad Y = \frac{z-\bar{z}}{i(1+|z|^2)}, \qquad Z = \frac{1-|z|^2}{1+|z|^2}. \tag{130}$$

In these coordinates, the symplectic form is

$$\omega = \frac{ik\, dz \wedge d\bar{z}}{(1+|z|^2)^2} \implies \omega_{z\bar{z}} = -\omega_{\bar{z}z} = \frac{ik}{(1+|z|^2)^2}, \tag{131}$$

where k is a constant. The normalization here can be determined from

$$\int_{\mathbb{R}^2} \omega = 2k \int \frac{1}{(1+r^2)^2} dx \wedge dy = 2\pi k \int_0^\infty \frac{2r dr}{(1+r^2)^2} = 2\pi k. \tag{132}$$

This normalization is chosen so that, if $k \in \mathbb{Z}$, the Berry connection θ is a canonically normalized $U(1)$ gauge field ($\int \omega \in \overline{\mathbb{Z}}$), meaning that each semiclassical state occupies a phase space volume of 2π. Note that $k \in \mathbb{Z}$ is required in order for integrals like $\oint dx^\mu \theta_\mu$ to make sense in an action (where the contour is around some curve in S^2), and so in what follows k will be an arbitrary integer. k determines the Chern class of the complex line bundle over $\mathscr{P}$, and as we will see later, determines how many states the quantum Hilbert space has.

It will be useful to know the derivatives of these coordinates wrt z and $\bar{z}$. Letting $\sigma \equiv 1 + |z|^2$, we find

$$\begin{aligned} dX &= \sigma^{-2}((1-\bar{z}^2)dz + (1-z^2)d\bar{z}), \qquad dY = -i\sigma^{-2}((1+\bar{z}^2)dz - (1+z^2)d\bar{z}), \\ dZ &= -\sigma^{-2}((1+\bar{z})dz + (1+z)d\bar{z}). \end{aligned} \tag{133}$$

One then checks that the Poisson brackets work out appropriately. For example,

$$\begin{aligned} \{X, Y\} &= \omega^{\mu\nu}\partial_\mu X \partial_\nu Y = -\frac{i\sigma^2}{k}(\partial X \bar{\partial} Y - \bar{\partial} X \partial Y) \\ &= -\frac{i}{k\sigma^2}\left[(1-\bar{z}^2)i(1+z^2) - (-i(1+\bar{z}^2)(1-z^2))\right] \\ &= 2\frac{1-|z|^4}{k\sigma^2} = \frac{2}{k}Z, \end{aligned} \tag{134}$$

and in general,

$$\{\widetilde{X}^i, \widetilde{X}^j\} = \epsilon^{ijk}\widetilde{X}^k, \qquad \widetilde{X}^i = \frac{k}{2}X^i, \tag{135}$$

which when quantized will give us the angular momentum algebra.

We now need to identify what the images of the X^i are under quantization. From the above calculation, we require that they satisfy

$$[\mathcal{Q}[X^i], \mathcal{Q}[X^j]] = i\hbar\epsilon^{ijk}\mathcal{Q}[X^k], \tag{136}$$

since the quantization map $\mathcal{Q}$ is linear (sorry for the minus sign! Too late to switch conventions).

First let us identify the Hamiltonian vector fields corresponding to each of the coordinates. This is done straightforwardly using

$$V^\mu_{X_i} = \omega^{\mu\nu}\partial_\nu X_i. \tag{137}$$

Some algebra gives

$$V_X = -\frac{i}{k}(1-z^2)\partial + \frac{i}{k}(1-\bar{z}^2)\bar{\partial}, \qquad V_Y = \frac{1}{k}(1+z^2)\partial + \frac{1}{k}(1+\bar{z}^2)\bar{\partial} \qquad V_Z = \frac{2i}{k}(z\partial - \bar{z}\bar{\partial}) \tag{138}$$

Since $\partial^\dagger = -\bar{\partial}$, all the vector fields are anti-Hermitian as differential operators.

We will also need the symplectic potential. We can take it to be either of the following choices:

$$\theta = \frac{ik\, z d\bar{z}}{1+|z|^2}, \qquad \text{or} \qquad \theta' = \frac{ik}{2(1+|z|^2)}(z d\bar{z} - \bar{z} dz). \tag{139}$$

When we contract with the various Hamiltonian vector fields, we get

$$\theta(V_X) = \sigma^{-1}(\bar{z}|z|^2 - z), \qquad \theta(V_Y) = i\sigma^{-1}(\bar{z}|z|^2 + z), \qquad \theta(V_Z) = 2\sigma^{-1}|z|^2. \tag{140}$$

Similarly,

$$\theta'(V_X) = -\frac{1}{2}(1-|z|^2)X, \qquad \theta'(V_Y) = -\frac{1}{2}(1-|z|^2)Y, \qquad \theta'(V_Z) = i(Z-1). \tag{141}$$

Using these vector fields, we can now write,[15] after some algebra,

$$\begin{aligned}
\mathcal{Q}'[X] &= \frac{\hbar}{k}(-(1-z^2)\partial + (1-\bar{z}^2)\bar{\partial}) + \frac{z+\bar{z}}{2}, \\
\mathcal{Q}'[Y] &= -i\frac{\hbar}{k}((1+z^2)\partial + (1+\bar{z}^2)\bar{\partial}) + \frac{z-\bar{z}}{2i} \\
\mathcal{Q}'[Z] &= \hbar(z\partial - \bar{z}\bar{\partial}) + 1.
\end{aligned} \tag{144}$$

[15]Note for posterity's sake: the following choices

$$\begin{aligned}
\mathcal{Q}[X] &= \frac{\hbar}{2}(-(1-z^2)\partial + (1-\bar{z}^2)\bar{\partial}) + \frac{z+\bar{z}}{1+|z|^2}, \\
\mathcal{Q}[Y] &= -i\frac{\hbar}{2}((1+z^2)\partial + (1+\bar{z}^2)\bar{\partial}) - i\frac{z-\bar{z}}{1+|z|^2} \\
\mathcal{Q}[Z] &= \hbar(z\partial - \bar{z}\bar{\partial}) + \frac{1-|z|^2}{1+|z|^2}.
\end{aligned} \tag{142}$$

also satisfy the correct commutation relations. These are the operators you get when you neglect the symplectic potential and write

$$\mathcal{Q}[f] = -i\hbar\nabla_{V_f} + f. \tag{143}$$

(also, here ω is normalized so that $\int\omega \in 4\pi\mathbb{Z}$). What's up? Is this a coincidence that this also works?

Or, if we use the other symplectic potential,

$$\begin{aligned}\mathcal{Q}'[X] &= \frac{\hbar}{k}(-(1-z^2)\partial + (1-\bar{z}^2)\bar{\partial}) + \bar{z},\\ \mathcal{Q}'[Y] &= -i\frac{\hbar}{k}((1+z^2)\partial + (1+\bar{z}^2)\bar{\partial}) + i\bar{z}\\ \mathcal{Q}'[Z] &= \frac{2\hbar}{k}(z\partial - \bar{z}\bar{\partial}) + 1.\end{aligned} \tag{145}$$

The operators which satisfy the usual angular momentum algebra are then

$$J_i \equiv \frac{k}{2}\mathcal{Q}[X_i]. \tag{146}$$

One can check that these satisfy the correct angular momentum algebra; I won't write out the algebra here.

Hilbert space

Having figured out the algebra, let's now figure out the Hilbert space. Like most of geometric quantization, this seems to be more an art than a science. Basically, the space of all functions on $\mathscr{P}$ is too big, so we need to choose a polarization to break it down to get the physical quantum Hilbert space. This is usually done by taking only those sections ψ that satisfy $i_X D\psi = 0$, where X is a section of some n-dimensional subbundle P of the full $2n$-dimensional bundle $L \to \mathscr{P}$. For example, in the case where $\mathscr{P}$ is a cotangent space, we might choose P to be the subbundle consisting of all the momenta, so that the sections remaining in the physical Hilbert space would be functions of coordinates only. In the Kahler case, a natural polarization is to choose P to be the subbundle of anti-holomorphic sections, so that the physical Hilbert space contains only the holomorphic sections, $\mathcal{H} = \mathrm{Hol}(L)$.

After we have chosen a polarization, we need to choose a metric—$\mathcal{H}$ then consists of the square-integrable sections ψ such that $i_X D\psi = 0$. This is sometimes a bit tricky—for example, in the case of $\mathscr{P} = \mathbb{R}^2$, if we were to work with the natural momentum polarization mentioned above, we would find no square-integrable sections, since all inner products of such polarized sections would contain the divergent integral $\int_{\mathbb{R}} dp$.

For cases where the phase space $\mathscr{P}$ is a Kahler manifold however, there is a natural inner product: choosing the polarization so that $\mathcal{H}$ consists of holomorphic sections of the complex line bundle $L \to \mathscr{P}$, the inner product is

$$\langle \psi, \eta \rangle = \int \omega\, e^{-\mathcal{K}} \psi^*(z)\eta(z), \qquad \psi, \eta \in \mathrm{Hol}(L). \tag{147}$$

Here $\mathcal{K}$ is the Kahler potential, which in the present setting of $\mathscr{P} = S^2$ is (note the factor of k; most references do not put the k here)

$$\mathcal{K} = k\ln(1+|z|^2). \tag{148}$$

In general the Kahler potential is such that $i\partial\bar{\partial}\mathcal{K} = \omega$. Indeed, we have

$$\theta = \frac{ik\bar{z}dz}{1+|z|^2} = i\partial\mathcal{K}, \tag{149}$$

which gives $i\partial\bar{\partial}\mathcal{K} = \omega$ since $d\theta = i\bar{\partial}\partial\mathcal{K} = \omega$. Note that the other choices of symplectic potential, like $i\bar{\partial}\mathcal{K}$, also give the correct symplectic form via $\omega = i\partial\bar{\partial}\mathcal{K}$, but the above choice is preferred, since with this choice $\theta_{\bar{z}} = 0$, meaning that we can define our holomorphic sections by the requirement $D_{\bar{X}}\psi = 0$, where $\bar{X} = \bar{\partial}$.[16]

Let's explain where the $e^{-\mathcal{K}}$ in the inner product comes from. While it initially looks rather mysterious, it actually just comes from requiring the integrand be well-defined when we change between patches on $\mathscr{P}$. Now in general, the local expression for a section ψ_α on different patches U_α, U_β, is

$$\psi_\alpha(z_\alpha) = g_{\alpha\beta}^k \psi_\beta(z_\beta), \tag{150}$$

where $g_{\alpha\beta}$ is the transition function on the line bundle L_1 with Chern class 1 (the bundle that the ψs are sections of is $L = L_1^{\otimes k}$), and is a holomorphic function of z_β (or z_α). Therefore

$$(\psi^*\eta)_\alpha = |g_{\alpha\beta}|^{2k}(\psi^*\eta)_\beta, \tag{151}$$

and so this part of the integrand is not well-defined. Therefore if (147) is to be well-defined, we need to have a compensating $|g_{\alpha\beta}|^{-2k}$ coming from the transformation of $\omega e^{-\mathcal{K}}$.

Indeed, such a compensation comes from the relations between the Kahler potentials on overlapping patches:

$$\mathcal{K}_\beta = \mathcal{K}_\alpha + k\left(\ln g_{\alpha\beta} + \ln g_{\alpha\beta}^*\right). \tag{152}$$

We see that this transformation rule leaves the symplectic form invariant and provides exactly the factors needed to ensure that the integrand is well-defined (i.e. independent of the choice of patch we express it in).

Let's check that this is indeed how $\mathcal{K}$ transforms for the case of $\mathscr{P} = \mathbb{CP}^N$. We will parametrize $\mathbb{CP}^N$ by $\{z_1, \ldots, z_{N+1}\}/\sim$, where $\sim$ is the usual rescaling relation. The patch U_i is defined to be the subset of $\mathbb{C}^{N+1}$ where $z_i \neq 0$, and we equip it with the coordinates $x_i^j = z_j/z_i$. The Kahler potential on U_i is

$$\mathcal{K}_i = k \ln\left[\sum_{j=1}^{N+1} |x_i^j|^2\right], \tag{153}$$

which reduces to the one written above for S^2 when $N = 1$. Now we see that the transition functions g_{ij} are x_i^j, since $x_i^l = x_i^j x_j^l = g_{ij} x_j^l$. Therefore we may write

$$\mathcal{K}_j = k \ln\left[|x_j^i|^2 \sum_{l=1}^{N+1} |x_i^l|^2\right] = \mathcal{K}_i + k(\ln g_{ji} + \ln g_{ji}^*), \tag{154}$$

which shows that $\mathcal{K}_i$ transforms between patches in the way that we claimed[17].

[16] With this choice, $D_{\bar{X}} = -i\hbar\bar{\partial}$; if we had let $\theta = i\bar{\partial}\mathcal{K}$ then we would have had $D_{\bar{X}} = -i\hbar\bar{\partial} + ikz/(1+|z|^2)$, and so we wouldn't be able to define holomorphic sections by the polarization $D_{\bar{X}}\psi = 0$.

[17] On S^2, sometimes different coordinates are preferred. Instead of working with $\zeta_N = (z_1/z_2, 1)$ and $\zeta_N = (1, z_2/z_1)$, we choose single-component coordinates $z_N = z, z_S = 1/z$. In this case, the transition function is $z_N = g_{NS} z_S$, with $g_{NS} = z_S^{-2}$. However, working with the ζ_Ns is preferred since although they have redundant components (since one of their components is always 1), their definition works for all $\mathbb{CP}^N$.

Now, what is the Hilbert space? Suppose that the largest power of z appearing in $\psi(z)$ is z^l. Then the norm of ψ is

$$||\psi|| \sim \int_{\mathbb{R}^2} dz \wedge d\bar{z} \frac{z^{2l}}{(1+|z|^2)^{2+k}} \sim \int_{\mathbb{R}} dr \, \frac{r^{1+2l}}{(1+r^2)^{2+k}}. \tag{155}$$

Therefore[18]

$$||\psi|| < \infty \implies 2+2l < 2(2+k) \implies l < k+1. \tag{156}$$

Therefore normalizable sections can be written as linear combinations of $1, z, \ldots z^k$, giving us $k+1$ different basis sections. This tells us that, for $j = k/2$,

$$\dim \mathcal{H} = k+1 = 2j+1, \tag{157}$$

which of course we could have guessed long ago. Therefore the result of geometrically quantizing sections of a $\mathbb{C}$ line bundle over S^2 with Chern number k is the theory of a quantum spin $k/2$ particle. Note that the number of states in the Hilbert space is one greater than the number we would have expected classically: $\dim \mathcal{H} = \int_{\mathscr{P}} \omega + 1$, although as expected the exact result agrees with the classical one when $k \sim \hbar^{-1} \to \infty$. The extra state in $\mathcal{H}$ in this case comes from the fact that $\mathscr{P}$ is positively curved: in general, positive phase-space curvature means $\dim \mathcal{H}$ is bigger than the semiclassical result, while negative curvature means $\dim \mathcal{H}$ is less. More on this in a future diary entry.

June 6 — Why exactly do Poisson brackets map to commutators during quantization?

Today we'll motivate why Poisson brackets become commutators under quantization, which I've only ever seen presented in a rather axiomatic way. We'll also prove some results needed in yesterday's diary entry.

Solution:

Before beginning, we will prove a math fact that will be useful later. We claim that for any two vector fields X, Y,

$$i_{[X,Y]}\Lambda = (\mathcal{L}_X i_Y - i_Y \mathcal{L}_X)\Lambda, \tag{158}$$

where $\Lambda \in \Omega^n(M)$ is any differential form on M, i_X is the interior product, and where $\mathcal{L}_X$ is the Lie derivative. My proof is inelegant, but oh well. First, we calculate

$$\mathcal{L}_X i_Y \Lambda = ([i_{\nabla_X Y}\Lambda]_{\boldsymbol{\lambda}} + [i_Y(\nabla_X \Lambda)]_{\boldsymbol{\lambda}} + Y^\mu \Lambda_{\mu\alpha\lambda_3\ldots\lambda_n} \partial_{\lambda_2} X^\alpha + \ldots) dx^{\boldsymbol{\lambda}}, \tag{159}$$

where $\boldsymbol{\lambda}$ is a multi-index $\boldsymbol{\lambda} = \lambda_2, \lambda_3, \ldots, \lambda_n$, so that $dx^{\boldsymbol{\lambda}} = dx^{\lambda_2} \wedge \cdots \wedge dx^{\lambda_n}$. Here the $+\ldots$ refers to other terms in the Lie derivative of Λ, viz. other indices that get contracted with

[18] We can also get this result without analyzing the integral by recalling that the transition function between the north and south patch means that z^l in the N patch gets mapped to z^{k-l} in the S patch. Now in order to not have a pole in the south patch we need $k-l \geq 0$, giving $l < k+1$ as above.

the derivative of X. I know the notation is opaque, but thinking of something better seemed to not be worth the trouble—if in doubt, just re-write it out. Anyway, the other relevant term is

$$\begin{aligned} i_Y \mathcal{L}_X \Lambda &= i_Y(\nabla_X \Lambda + \partial_{\lambda_1} X^\mu \Lambda_{\mu \boldsymbol{\lambda}} dx^1 \wedge dx^{\boldsymbol{\lambda}}) \\ &= i_Y(\nabla_X \Lambda) + i_{\nabla_Y X} \Lambda - \partial_{\lambda_2} X^\mu Y^\alpha \Lambda_{\mu \alpha \lambda_3 \ldots \lambda_n} dx^{\boldsymbol{\lambda}}. \end{aligned} \tag{160}$$

We can flip the sign on the last term by exchanging the μ and α indices in Λ, and when this is subtracted from (159), it will cancel with the last term there. Similarly, the first term in the above equation will cancel with the second in (159). This leaves only two terms, and we find that indeed,

$$(\mathcal{L}_X i_Y - i_Y \mathcal{L}_X)\Lambda = (i_{\nabla_X Y} - i_{\nabla_Y X})\Lambda = i_{[X,Y]}\Lambda. \tag{161}$$

Now we can talk about Poisson brackets. Let X_f and Y_g be two Hamiltonian vector fields on phase space associated to Hamiltonians f and g, respectively, and let ω be the symplectic form. This means that

$$i_{X_f}\omega = -df \implies \omega_{\mu\nu} X_f^\mu = \partial_\mu f, \qquad X_f^\mu = \omega^{\mu\nu}\partial_\nu f, \tag{162}$$

and likewise for Y_g. Keeping track of the signs here is supremely important, unfortunately. We have made a choice that means commutators will become negative Poisson brackets, but this was only realized retrospectively—oh well, can't be helped now.

Our goal will be to related the vector field $[X_f, Y_g]$ to the vector field $V^\mu_{\{f,g\}} = \omega^{\mu\nu}\partial_\nu\{f,g\}$ (if f, g are Hamiltonians, then their Poisson bracket also generates a Hamiltonian vector field). In the following, we will drop the subscripts on X and Y for clarity.

First, we calculate, using the result derived above,

$$i_{[X,Y]}\omega = (\mathcal{L}_X i_Y - i_Y \mathcal{L}_X)\omega = \mathcal{L}_X i_Y \omega = \mathcal{L}_X(Y^\sigma \omega_{\sigma\nu})dx^\nu = -\mathcal{L}_X dg, \tag{163}$$

where our sign convention is such that $Y^\sigma \omega_{\sigma\nu} = -\partial_\nu g$, and where we used the stationarity of ω under the flow of X:

$$\mathcal{L}_X \omega = (di_X + i_X d)\omega = d(i_X\omega) = d(-df) = 0. \tag{164}$$

Expanding this out, we have

$$\begin{aligned} -[\mathcal{L}_X dg]_\nu &= -X^\mu \partial_\mu \partial_\nu g - \partial_\nu X^\mu \partial_\mu g = -\omega^{\mu\lambda}\partial_\lambda f \partial_\mu \partial_\nu g - \partial_\nu(\omega^{\mu\lambda}\partial_\lambda f)\partial_\mu g \\ &= \partial_\nu(\omega^{\lambda\mu}\partial_\lambda f \partial_\mu g) = \partial_\nu\{f,g\}. \end{aligned} \tag{165}$$

Now the vector field generated by the Poisson bracket is

$$V^\mu_{\{f,g\}} = \partial_\lambda(\omega^{\sigma\nu}\partial_\sigma f \partial_\nu g)\omega^{\mu\lambda}, \tag{166}$$

so that

$$i_{V_{\{f,g\}}}\omega = -\partial_\nu(\omega^{\sigma\lambda}\partial_\sigma f \partial_\lambda g)dx^\nu = -\partial_\nu\{f,g\}dx^\nu, \tag{167}$$

which when compared with (165) and (163) gives

$$i_{[X,Y]}\omega = -i_{V_{\{f,g\}}}\omega. \tag{168}$$

Since ω is invertible, the two vector fields in the interior products must be negatives of each other, and hence (restoring subscripts on X and Y for clarity)

$$[X_f, Y_g] = -V_{\{f,g\}}. \tag{169}$$

Graahh, that minus sign is annoying. Oh well.

This tells us that the commutator of Hamtilonian vector fields, when represented as differential operators, generates the Poisson bracket algebra. In geometric quantization we send functions on phase space (f and g) to differential operators (first order ones in e.g. Kahler quantization) that perform the flow along their respective vector fields. Therefore if we very crudely think of (pre)quantization as a mapping of functions on phase space to differential operators via $f \mapsto X_f$, the above equation tells us that the quantization of the Poisson bracket $\{f, g\}$ is the commutator of the quantizations of f and g.

June 7 — Scale invariance, conformal invariance, and algebras of charge operators

Today we're going to do an elaboration on an exercise given by Slava Rychkov during his 2019 TASI lectures.

Solution:

First, we will only assume translation and rotation symmetry. We want to calculate $Q_\epsilon(\Sigma) \cdot T^{\mu\nu}(0)$, for Σ a surface enclosing 0, which we can do by general reasoning. The only two ϵs we have access to for translations and rotations are $\epsilon_\mu = a_\mu$ with a_μ constant (translations) and $\epsilon_\mu = \omega_{\mu\nu}x^\nu$ with $\omega_{\mu\nu}$ constant and antisymmetric (rotations). Now since $T^{\mu\nu}$ is conserved, the Ward identity tells us that its dimension is exactly 2. Since $Q_\epsilon(\Sigma) \cdot T$ needs to be linear in ϵ, dimensional analysis then tells us that each term appearing in $Q_\epsilon(\Sigma) \cdot T$ should contain one ϵ, one T, and one ∂. Therefore we may write

$$Q_\epsilon(\Sigma) \cdot T_{\mu\nu} = \epsilon^\rho \partial_\rho T_{\mu\nu} + A\partial_\mu \epsilon^\rho T_{\rho\nu} + B\partial_\nu \epsilon^\rho T_{\rho\mu} \tag{170}$$

for some constants A, B. The coefficient of the first term has been fixed to one because we want momentum to act on all operators as $P_\epsilon = \epsilon^\mu \partial_\mu$. By $\mu \leftrightarrow \nu$ symmetry, we have $B = A$, and taking the divergence of this then tells us that $A = 1$.[19] This means we can write

$$Q_\epsilon(\Sigma) \cdot T_{\mu\nu} = \epsilon^\rho \partial_\rho T_{\mu\nu} - \partial^\rho \epsilon_\mu T_{\rho\nu} + \partial_\nu \epsilon^\rho T_{\rho\mu}, \tag{171}$$

where we've used the Killing equation for ϵ in the second term to change the sign.

Now we want to ask what algebra the charges obey. The commutator of charges is computed in the usual way:

$$[Q_\epsilon(\Sigma), Q_\eta(\Sigma)] = \int_\Sigma d^{d-1}x^\mu \int_{S_x^{d-1}} d^{d-1}y^\alpha \, T_{\mu\nu}(x)\eta^\nu(x)T_{\alpha\beta}(y)\epsilon^\beta(y), \tag{172}$$

[19] Use $\partial_\mu \partial^\mu \epsilon^\rho = -\partial_\mu \partial^\rho \epsilon^\mu = -\partial^\rho(\partial \cdot \epsilon) = 0$.

where S_x^{d-1} is a sphere centered on x. Shrinking down this sphere and then using (171), we have

$$[Q_\epsilon(\Sigma), Q_\eta(\Sigma)] = \int_\Sigma d^{d-1}x^\mu \, \eta^\nu(\epsilon^\rho \partial_\rho T_{\mu\nu} - \partial^\rho \epsilon_\mu T_{\rho\nu} + \partial_\nu \epsilon^\rho T_{\rho\mu}). \tag{173}$$

The annoying thing about this is that the derivatives are derivatives in the full $\mathbb{R}^d$, and not covariant derivatives for the metric restricted to the surface Σ. This makes integrating by parts annoying if Σ is chosen to be some generic curved surface. However, since the charge operators are topological, we can choose any representative choice of Σ that we like—in order for integration by parts to be done easily, we will therefore choose Σ to be a hypercube. Integrations by parts can then be performed on each cube face, with the boundary terms all canceling out pairwise. Let us focus on the face F_1 of the cube with unit normal in the x^1 direction, just for concreteness. Then the first integral in (173) is

$$\int_{F_1} d^{d-1}x^1 \, \left[\eta^\nu \epsilon_1 \partial^1 T_{1\nu} - \epsilon_\rho \partial^\rho \eta^\nu T_{1\nu} + \epsilon_1 \partial^1 \eta^\nu T_{1\nu}\right] + \ldots, \tag{174}$$

where $\ldots$ is a boundary term coming from integrating by parts on F_1—these terms will end up canceling when we sum over cubes, so we will avoid writing them explicitly. The first and last terms above can be combined as

$$\int_{F_1} d^{d-1}x^1 \, \left[\partial^1(\eta^\nu \epsilon_1 T_{1\nu}) - (\nabla_\epsilon \eta)^\nu T_{1\nu}\right], \tag{175}$$

since ϵ is Killing (here $\nabla_\epsilon = \epsilon^\rho \partial_\rho$).

The next term we will deal with is the second term in (173). We can integrate by parts and write this as

$$-\int_{F_1} d^{d-1}x^1 \, (\partial_1 \eta^\nu \epsilon_1 T_{1\nu} + \eta^\nu \partial^1(\epsilon_1 T_{1\nu}) - \partial_\rho \eta^\nu \epsilon_1 T_{\rho\nu}). \tag{176}$$

By the symmetry of $T_{\mu\nu}$ and the Killing equation for η, the last term vanishes, and so this becomes

$$-\int_{F_1} d^{d-1}x^1 \, \partial_1(\eta^\nu \epsilon_1 T_{1\nu}). \tag{177}$$

Note that this is precisely the right term to cancel the total ∂^1 derivative in (175). Therefore there are only two terms in (173) which survive, which together give

$$[Q_\epsilon(\Sigma), Q_\eta(\Sigma)] = \int_\Sigma d^{d-1}x^\mu \, T_{\mu\nu}[(\nabla_\eta \epsilon)^\nu - (\nabla_\epsilon \eta)^\nu] = Q_{[\eta,\epsilon]}(\Sigma). \tag{178}$$

Note the "backwards" order on the Lie bracket!

Now let us add the possibility that the theory also possesses scale invariance: therefore we can allow for transformations with $\partial_{(\mu}\epsilon_{\nu)} \propto \delta_{\mu\nu}$, provided that $\mathrm{Tr}[T] = 0$. When we put this condition on T, we can allow a new term in the $Q_\epsilon \cdot T_{\mu\nu}$ OPE, viz.

$$Q_\epsilon(\Sigma) \cdot T_{\mu\nu} = \epsilon^\rho \partial_\rho T_{\mu\nu} - \partial^\rho \epsilon_\mu T_{\rho\nu} + \partial_\nu \epsilon^\rho T_{\rho\mu} + \partial_\rho \epsilon^\rho T_{\mu\nu}, \tag{179}$$

where the coefficient of the last term is fixed by requiring that the RHS be conserved. Now when we compute the charge commutator, we get, on each face, the integral

$$[Q_\epsilon(\Sigma), Q_\eta(\Sigma)] \ni \int_{F_1} d^{d-1}x^1\, \eta^\nu [\epsilon^\rho \partial_\rho T_{1\nu} - \partial^\rho \epsilon_1 T_{\rho\nu} + \partial_\nu \epsilon^\rho T_{\rho 1} + \partial_\rho \epsilon^\rho T_{1\nu}]. \tag{180}$$

Integrating by parts and dropping the boundary terms that will cancel, the first term in the above is

$$\int_{F_1} d^{d-1}x^1 \left[-\nabla_\epsilon \eta^\nu T_{1\nu} - \eta^\nu (\partial\cdot\epsilon) T_{1\nu} + \partial_1(\eta^\nu \epsilon^1 T_{1\nu})\right]. \tag{181}$$

The first term contains the part of the Lie bracket that we want, while the second term cancels with the third term in (180). Therefore we get

$$[Q_\epsilon(\Sigma), Q_\eta(\Sigma)] \ni \int_{F_1} d^{d-1}x^1 \left[(\nabla_\eta \epsilon^\nu - \nabla_\epsilon \eta^\nu) T_{1\nu} + \partial_1(\eta^\nu \epsilon^1 T_{1\nu}) - \eta^\nu \partial^\rho \epsilon_1 T_{\rho\nu}\right]. \tag{182}$$

The last term is however

$$-\int_{F_1} d^{d-1}x^1 \left[\partial_1(\eta^\nu \epsilon_1 T_{1\nu}) - \epsilon_1 \partial_\rho \eta^\nu T_{\rho\nu}\right]. \tag{183}$$

The first term cancels the second term in (182), while the second term vanishes for the usual reasons of symmetry and tracelessness of T. Therefore we again get $[Q_\epsilon(\Sigma), Q_\eta(\Sigma)] = Q_{[\eta,\epsilon]}(\Sigma)$.

Finally, some comments on why $d = 2$ is special (thanks to the Havard CFT guys at TASI for helping me understand this).

We can take ϵ to be *any* holomorphic function. Since the inner product on the plane is..., only functions with simple poles (at zero or infinity) give global charges. That is, the only global charges are those corresponding to $\epsilon = z, z^{-1}, c$ with c a constant.

This singularity is okay, provided that it exactly coincides with the insertion point of an operator. When we normally discuss the algebra of the L_{-n}s, we talk about their action on states, i.e. on operators located at the origin on the plane.

June 9 — Free Maxwell is not conformally invariant in $d \neq 4$

Today I'm going to write down some stuff I learned that was inspired during questions in one of Slava Rychkov's 2019 TASI lectures. We will see that free Maxwell theory, while scale invariant in any dimension, is only conformally invariant in $d = 4$. A good chunk of our discussion is going through stuff in Rychkov's earlier paper [**?**].

Solution:

F is not a primary because its 2-point functions don't have the right tensor structure

Let

$$\Gamma_{\mu\nu} \equiv g_{\mu\nu} - d\frac{x_\mu x_\nu}{x^2}. \tag{184}$$

Then

$$\langle F_{\mu\nu} F_{\lambda\sigma} \rangle = \frac{d-2}{x^d} \left(\Gamma_{\mu\lambda} g_{\nu\sigma} - \Gamma_{\nu\lambda} g_{\mu\sigma} - \Gamma_{\mu\sigma} g_{\nu\lambda} + \Gamma_{\nu\sigma} g_{\mu\lambda} \right). \tag{185}$$

Actually, we could have written down all of this correlator without doing any calculations (though we wouldn't be able to fix the overall prefactor). Since the theory is free and F is conserved, the dimension of the expression is fixed. Since the F operators are normal-ordered, the components in the tensor structure will never pair up μ with ν or λ with σ; however, the correlator must be antisymmetric under the exchange of both pairs of indices. Finally, the prefactor $(d-2)$ can be fixed since we know that in $d = 2$ there are no propagating degrees of freedom.

F is not a primary because it doesn't transform like one under conformal transforamtions

We now give a different proof for why $F_{\mu\nu}$ isn't a primary, by looking at the way it transforms under conformal transformations.

A primary field transforms under finite conformal transformations as $\Phi(x) \mapsto \Phi'(x')$, where

$$\Phi'(x') = \Omega(x)^{-\Delta} R(x) \cdot \Phi(x), \tag{186}$$

where our notation $R \cdot \Phi$ means the action of the rotation R in whatever representation of $SO(d)$ Φ is valued in. Infinitesimally, for $x' = x + \xi$ a transformation determined by the vector field ξ, we have

$$\delta_\xi \tag{187}$$

Now let us look at how the linear combinations $C^{i\mu}\partial_\mu \Phi_i$ transform. Our goal will be to find out when such linear combinations can be primaries.

June 8 — Rant about misunderstandings of "diffeomorphism invariance" in physics

The vast majority of statements made (by physicists) regarding the role of diffeomorphisms in physics are imprecise / wrong, which has caused me a lot of confusion recently.

Now the abuse of math terminology is itself not a big problem, provided that one abuses terminology in a consistent way which doesn't interfere with physical understanding. In my experience though, the sloppy use of mathematical terminology tends to lead people (like me) to get confused about the physics, which is no good. So, today we will try to set things straight and provide an unambiguous definition for all the terms that usually get thrown about haphazardly.

Solution:

Note: we will be working with a level of rigor halfway between that of a physicist and a mathematician. The goal is to provide enough precision to do away with the ambiguities present in most physics discussions, while staying away from being overly fussy about notation.

Let us recall some terminology. For us, a manifold will be a smooth, differentiable manifold. So, for us, a manifold M is a Hausdorff space[20] which is locally homeomorphic[21] to $\mathbb{R}^n$. The identification with $\mathbb{R}^n$ is accomplished locally by the coordinate functions ϕ_a, which map open subsets U_a of M (we will assume that these patches U_a are all simply connected) into $\mathbb{R}^n$ via $\phi_a : U_a \to \mathbb{R}^n$. The epithet "smooth, differentiable" just means that the transition functions $\phi_a \circ \phi_b^{-1}$, defined on the image of $\phi_a(U_a \cap U_b)$, are $\mathcal{C}^\infty$ functions.

Diffeomorphisms

A diffeomorphism is a smooth map between manifolds that is invertible and whose Jacobian matrix is invertible everywhere. That is, a diffeomorphism between two same-dimensional manifolds M and N is an invertible map f which assigns points in M to points in N, and which is such that if $\phi_{N/M}$ are coordinate maps, then the function $\phi_N \circ f \circ_M^{-1} : \mathbb{R}^{\dim M} \to \mathbb{R}^{\dim M}$ is differentible and has invertible Jacobian everywhere.[22] Basically, generically any "nice" thing one might imagine doing to a manifold is a diffeomorphism. Some potentially "not nice" things are also diffeomorphims. For example, the Dehn twist on one of the cycles of a T^2 is a diffeomorphism. It is a large diffeomorphism since it cannot be built out of smaller ones that are connected to the identity map: cutting the torus and rotating one of the cut boundaries by any angle $\theta \notin 2\pi\mathbb{Z}$ does not map neighboring points to neighboring points, and violates the smoothness condition.

The condition on the invertibility of the Jacobian ensures that the tangent spaces on M and N get mapped to eachother in a bijective way. Therefore e.g. $x \mapsto x^2$ is not a diffeomorphism since it isn't invertible, and $x \mapsto x^3$ isn't a diffeomorphism from $\mathbb{R}$ to itself for since although it is differentiable and invertible, the Jacobian fails to be invertible at zero.[23] More generally, the problem is that when f fails to have an intertible Jacobian, we can no longer pull back vectors in T_yN to vectors in $T_{f^{-1}(y)}M$, since if $f(x) = y$ then

$$(f^*V)^\mu(x) = \frac{\partial x^\mu}{\partial y^\alpha}\Big|_y V^\alpha(f(x)), \qquad V \in T_yN, \tag{188}$$

[20] A topological space is basically a set X (a bunch of points) together with a collection of all neighborhoods $N(x)$ of all points $x \in X$, with obvious consistency conditions ($x \in N(x)$ for all x, $N(x)$; if $N(x), N'(x)$ are two neighborhoods of x then $N \cap N'$ is also a neighborhood of x; every neighborhood $N(x)$ includes a sub-neighborhood $N'(x)$ such that $N(x)$ is a neighborhood of all points $y \in N'(x)$). A Hausdorff space is a topological space which is such that for any two distinct points $x, y \in N(x)$, there are always neighborhoods of x and y which are disjoint.

[21] A homeomorphism is a continuous + invertible map between two topological spaces. It is defined without reference to calculus, so it does not have to be "smooth" in any sense.

[22] We don't want to keep track of all the notational baggage that comes along with explicitly mentioning the coordinate functions of M and N, and so in what follows we will just write e.g. $f(x) = y$ for the diffeomorphism, where x, y are coordinates on M and N (images of points in M, N under coordinate maps), respectively.

[23] f is not injective on the tangent space, since the preimage of $T_0\mathbb{R}$ is not defined.

which doesn't make sense when the Jacobian of f isn't invertible.[24]

Another way to state this, if we want practice with the terminology, is that f being a diffeomorphism means that the pushforward f_*, which maps $TM \to TN$, has zero kernel everywhere. Recall that vectors in T_xM are pushed forward via

$$(f_*V)^\alpha(y) = \frac{\partial y^\alpha}{\partial x^\mu}\Big|_x V^\mu(f^{-1}(y)), \qquad V \in T_xM, \tag{190}$$

which has a kernel if Df isn't invertible.

So, a diffeomorphism $f : M \to N$ is essentially a smooth invertible map between) points on the two manifolds, which is such that it extends to a smooth invertible map between T_pM and $T_{f(p)}N$ for all p. Note that the definition of a diffeomorphism does *not* refer to any metric structure! A manifold is defined by a collection points and a collection of coordinate charts, and does not come equipped with any notion of distance. Indeed, let (M, g), (N, h) be two Riemannian manifolds. Then a diffeomorphism $f : M \to N$ is defined in exactly the same way as above, without reference to a metric (note that we write $f : M \to N$ and not $f : (M, g) \to (N, h)$!). For example, the unit disk with the flat Euclidean metric is diffeomorphic to the Poincare disk. This is a pretty trivial diffeomorphism, with a point on the flat disk mapped to the same point on the Poincare disk. This diffeomorphism doesn't care that it maps two points that are close together in the flat metric to two points that are (perhaps very) far apart on the Poincare disk—diffeomorphisms are smooth associations of points on one manifold to points on another, nothing more. In particular, diffeomorphisms do not generically preserve any sort of curvature. Therefore the notion of a diffeomorphism is a very broad one: two manifolds that are diffeomorphic are isomorphic with regards to differential topology: if one only cares about the smooth structure, then diffeomorphic manifolds are the same. Basically, all reasonable things that a physicist might imagine doing to a manifold are diffeomorphisms.

Sometimes one hears a big fuss being made about active and passive diffeomorphisms. As far as I can tell, an active diffeomorphism is a diffeomorphism $f : M \to M$ which moves points on M to other points on M, but keeps the coordinate charts unchanged. On the other hand, a passive diffeomorphism $f : M \to M$ is one which keeps points on the manifold fixed, but changes the coordinate functions ϕ_a. I don't really see the big deal of distinguishing between these two equivalent ways of thinking—sometimes passive diffeomorphisms are viewed as trivial "coordinate transformations" while active diffeomorphisms are viewed as something more deep that only GR possesses invariance under (see e.g. stuff by Rovelli). I don't see the point in distinguishing between them, since separating a manifold into a topological space ("a background spacetime") and "a coordinate system" and treating these two as if they

[24]Given any smooth map $f : M \to N$, we can always pull back fields on T^*N to fields on T^*M: for every $y = f(x)$, we can write

$$(f^*V)_\mu(y) = \frac{\partial y^\alpha}{\partial x^\mu}\Big|_x V_\alpha(x), \qquad V \in T^*N. \tag{189}$$

By contrast, vector fields can only be pulled back when we can "swap upper and lower indicies" in the above, that is, when Df is invertible—i.e. when f is a diffeomorphism. Likewise, for general smooth f, we can always push forward vector fields in TM to vectors fields in TN, but we cannot push forward fields in T^*M, unless f is a diffeomorphism.

were separate things is not useful—spacetime is the background topological space *and* the coordinate charts; both are needed to define a spacetime manifold.

Anyway, what would it mean for a theory to be "diffeomorphism invariant"? Such a theory would have to be completely insensitive to smooth changes in the manifold on which it is paced. Since diffeomorphisms do not care about distances, such a theory would need to be completely insensitive to sizes, shapes, and curvature. Of course, such a theory would therefore need to be topological. Gravity is most emphatically not diffeomorphism invariant—saying that "GR is diffeomorphism invariant" is saying that gravity is topological, which (at least in $d \geq 4$) is totally fatuous and completely antithetical to the spirit of GR.

Isometries

In my experience, physicists also almost always use the word "isometry" incorrectly. An isometry is what physicists often call a diffeomorphism. That is, an isometry is a diffeomorphism between Riemann manifolds $f : (M, g) \to (N, h)$ such that distances are preserved, viz. for any two vector fields v, w on M, we have[25]

$$g(V, W) = h(f_* V, f_* W), \qquad \forall\, V, W \in TM, \tag{191}$$

where in our notation, a metric g is defined as a section of $T^*M \times T^*M$, i.e. an assignment of a (0,2)-type tensor to each point of M. A simpler way of writing the above is to say that the metrics g, h are related by pullback along f:

$$g = f^* h. \tag{192}$$

Hence isometries are diffeomorphisms $f : (M, g) \to (N, f^*g)$. Just to make the notation vis-a-vis the metric totally clear,

$$g(v, w) = g_{\mu\nu} V^\mu W^\nu, \qquad h(f_* V, f_* W) = h_{\alpha\beta} \frac{\partial y^\alpha}{\partial x^\mu} \frac{\partial y^\beta}{\partial x^\nu} V^\mu W^\nu, \tag{193}$$

where the metric components are e.g.

$$g_{\mu\nu} = g(\partial_\mu, \partial_\nu). \tag{194}$$

So, in components, this means that under an isometry, the new metric h is determined from the old metric g by

$$h_{\alpha\beta}(y) = \left(\frac{\partial x^\mu}{\partial y^\alpha} \frac{\partial x^\mu}{\partial y^\alpha} \right) \Bigg|_x g_{\mu\nu}(x). \tag{195}$$

Note that for f to be an isometry, M and N do not have to be the same, they only need to be related by a diffeomorphism, with their metrics related as above. In particular, isometries are *not* transformations that preserve the components of the metric—they preserve distances, and calculating distances is done with more than just the components of the metric. A

[25] Note that since f is a diffeomorphism, $f_* V$ never vanishes unless V does, and so positive-definiteness is preserved.

very simple example: consider the rescaling diffeomorphism of $\mathbb{R}^n$ to itself given by $y = \lambda x$. This is an isometry provided that the metric maps as $g \mapsto h = \lambda^{-2} g$, since then while the metric components change, distances remain the same, since inner products map as $g(V, W) \mapsto \lambda^{-2} g(\lambda V, \lambda W) = g(V, W)$.

To sum up: isometries generically change both the coordinates and the metric, and are essentially geometrically trivial, since they do not change distances (or curvature). Just as a diffeomorphism is an isomorphism in the sense of differential topology, an isometry is an isomorphism in the sense of Riemannian geometry.

In my experience, physicists usually refer to isometries as "coordinate transformations" or "diffeomorphisms" or "isometries", and a theory which is invariant under isometries is often referred to as "covariant" or "diffeomorphism invariant" or "diffeomorphism covariant". It's a mess. The physicists who know what they are talking about and who refer to isometries as diffeomorphisms actually really aren't committing that big a sin, since given a diffeomorphism between a manifold M and itself, we can always construct an isometry $f : (M, g) \to (M, f^* g)$. When physicists mean "diffeomorphism", they usually mean a specific diffeomorphism, where the metric on the target manfiold is related by pullback to the original metric. I still think that it is better to call this an isometry, though.

When people say "physics can't depend on a change of coordinates", this is true as long as they tacitly mean "provided that the metric changes accordingly with the coordinates in such a way as to produce an isometry". Just changing the coordinates is certainly *not* a physically trivial operation, since if we don't touch the metric, then changing coordinates can change distances, curvature, etc. However, changing both the coordinates and the metric in such a way as to produce an isometry really is trivial, and it is the invariance of physical theories under this combined procedure that people generally refer to as "coordinate independence".

Spacetime symmetries in field theory

Now that we know what diffeomorphisms and isometries are, what then is a "spacetime symmetry" in physics parlance?

blah blah autoisometry

In field theory, symmetries act *only on the dynamical fields*. They do not act on the coordinates (which are just integration variables), and neither do they act on any background fields present. In non-gravitational theories, the metric is a background field; hence symmetries do not touch the metric. This fact led to a lot of confusion for me, since it's often convenient to think of the action of a spacetime symmetry as an action on the metric. We also are often in the situation where we'd like to compare the parition functions for theories with two different metrics, e.g. when we want to compare partition functions of a theory on spheres of different radii. But in some sense different choices of background metric lead to different theories: this is not so for symmetries, which always relate one theory to itself.

The clearest way to discuss spacetime symmetries is probably to look at correlation functions.

Now lets discuss actions.

It's conceptually to define vector fields on field space as

$$\mathcal{V}_\xi = \int d^d x \, \mathcal{L}_\xi \phi \frac{\delta}{\delta \phi}, \tag{196}$$

where ϕ^a stand for all the fields in the theory (the repeated a index means that we are to sum over all fields), and the $\delta/\delta\phi^a$ are basis vectors in the tangent space of field configurations. We will write the Lie derivative on field space as $\mathscr{L}$, so that

$$\mathscr{L}_\xi F = ... \tag{197}$$

for any function F of the fields.

We will let L denote the Lagrangian in the theory, viewed as a d-form. So e.g. we can have $L \sim F \wedge \star F$ or $L \sim d\phi \wedge \star d\phi - \phi^4 |d^d x|$, where $|d^d x|$ is the volume d-form. Writing L as a form is helpful since Cartan's formula tells us that

$$\mathcal{L}_\xi L = (i_\xi d + d i_\xi) L = d(i_\xi L), \tag{198}$$

where i_ξ is the contraction and we used $dL = 0$ on account of L being top-dimensional. Therefore, the Lie derivative of L along any vector field is always a total derivative. Our condition for a symmetry is that L be unchanged up to a total derivative, and so our condition for a symmetry is that

$$\mathcal{L}_\xi L = \mathscr{L}_\xi L. \tag{199}$$

If $f^* g = g$, then g must be constant along the flow of f, and so its Lie derivative must vanish

$$\mathcal{L}_\xi g = 0. \tag{200}$$

In components,

$$\begin{aligned} g_{ab}(x + \epsilon\xi) = g_{ab}(x) + \epsilon\big[(\xi \cdot \partial) g_{ab}(x) + \partial_a \xi^c g_{cb} + \partial_b \xi^c g_{ac}\big] + O(\epsilon^2) \\ \implies \mathcal{L}_\xi g_{ab} = (\xi \cdot \partial) g_{ab}(x) + \partial_a \xi^c g_{cb} + \partial_b \xi^c g_{ac}. \end{aligned} \tag{201}$$

Requiring that $\mathcal{L}_\xi g = 0$ implies, after some algebra needed to covariantize the derivatives above, that ξ satisfies the Killing equation, viz.

$$\nabla_{(\mu} \xi_{\nu)} = 0. \tag{202}$$

A symmetry can be defined in a situation where one has an isometry where both M and N, as well as g and h, happen to be the same. The fact that M and N are the same means that f maps points on M to points on M. Thus a symmetry is an autoisometry. Thus symmetries are what are often referred to as "isometries" in physics.

These are associated with Killing vectors: they preserve the form of the metric *components*, i.e. are such that $g = f^* g$: pulling the metric back along the diffeomorphism leaves it invariant. In a general isometry, the pulled-back metric is distinct from the original metric.

However, one note of caution: the statement that "changing the coordinates is equivalent to changing the metric" is only true in theories where isometries act trivially. For example, consider an action with a translationally non-invariant term like

$$S \ni \int d^d x \, \cos(x) \phi^2 \tag{203}$$

(even if this is messed up as a QFT, we can still think about this as occuring in some weird stat-phys context). This term is not translationally invariant, since (duh) when we change

coordinates by $x \mapsto x + a$, the action is changed. However, the metric doesn't change under a translation, and so changing the coordinates is not the same as changing the metric: one changes the action, and one doesn't.

If the metric is unchanged under translations, why does T_{0i} generate translations, given that it can be derived by varying wrt the metric?

philosophy about only knowing about symmetries when you can break them and ward ids

GR

The discussions concerning GR are usually most misleading—this is very annoying, because even ostensibly reputable sources get the discussion of diffeomorphisms wrong. The best discussion I ended up finding seemed to be in Carroll. Although his use of the word "diffeomorphism" is technically speaking mathematically not correct, if we replace his "diffeomorphism" with "isometry", his discussion is pretty much spot-on.

Anyway, from the above discussion, we can see that the shibboleth "GR is special because it is diffeomorphism invariant" really ought to be done away with. If GR were diffeomorphism invariant, it wouldn't be a geometric theory at all! When people call GR diffeomorphiism invariant, they really mean either 1) GR is invariant under isometries or 2) they mean the equations of GR are covariant (made out of tensors) which is a misleading + contentless statement since this is true for all fundamental physical theories or 3) they are confused.

The field equations $G_{\mu\nu} = 8\pi T_{\mu\nu}$ are invariant under pulling back both sides

The thing that distunguishes GR from other theories is that now it is the metric we are solving for. If we take the physicists approach of thinking of the metric in terms of its components, then the fact that the field equations are invariant under isometries .... So there is a redundancy in the metric components. But if we instead think of the metric as a section of $T * M \times T * M$, then there is no redundancy; it is strictly invariant under isometries.

Conformal transformations

Now we come to another conceptual mess: the exact definition of a conformal transformation in QFT. After asking O(10) people who do way more CFT than I do, I got a pretty much random smattering of answers to "do conformal transformations change curvature?" and "are correlation functions invariant or covariant under conformal transformations". Here we'll try to get everything straight.

Conformal transformations are diffeomorphisms, but they are not all isometries. A conformal transformation is a diffeomorphism $f : (M, g) \to (N, h)$ such that the pullback of the metric on N is related to the metric on M up a Weyl rescaling factor:

$$f^* h = \Omega g. \tag{204}$$

In components, this reads

$$h_{\alpha\beta}(y) = \Omega(x) \left(\frac{\partial x^\mu}{\partial y^\alpha} \frac{\partial x^\mu}{\partial y^\alpha} \right) \Bigg|_x g_{\mu\nu}(x). \tag{205}$$

blah blah

$$\mathcal{L}_\xi g = \omega g, \tag{206}$$

which, when written out in terms of covariant derivatives, is

$$\nabla_{(\mu}\xi_{\nu)} = \omega g_{\mu\nu}. \tag{207}$$

Another thing we should get straight is the notion of a Weyl transformation. A Weyl transformation . Thus a conformal transforamtion is a Weyl transformation, but the converse is not necessarily true.

June 14 — Conservation of $T^{\mu\nu}$, different kinds of stress tensors, and useful covariant derivative identities

Today we're doing a short calculation showing why $T^{\mu\nu}$ is conserved. There's a nontrivial integration by parts that is normally glossed over which we will explain. We'll also explain the relation of the two different usual ways of calculating $T_{\mu\nu}$ (by varying the fields or by varying the metric), which is annoyingly not really explained correctly in any of my books.

Solution:

The covariant conservation of the stress tensor in a given theory follows from the invariance of the action under isometries (often called diffeomorphisms by physicists; see yesterday's entry).

Consider an infinitesimal flow along a diffeomorphism $f : M \to M$, where the coordinates are mapped as $x^\mu \mapsto x^\mu + \xi^\mu$. The Lagrangian $d^d x\, \sqrt{g}\, \mathcal{L}$ is unchanged under both mapping the coordinates under the diffeomorphism and pulling back the fields along the diffeomorphism, since the combination of both is a reparametrization of our coordinate system, under which all theories whose Lagrangians don't explicitly depend on the coordinates (i.e. all theories in physics) are invariant.

After mapping the fields under the diffeomorphism by replacing them with their pullbacks (but not changing the coordinates), the Lagrangian changes as

$$\delta(\sqrt{g}\mathcal{L}) = \left(\mathcal{L}_\xi \phi \frac{\delta}{\delta\phi} + (\mathcal{L}_\xi g)_{\mu\nu}\frac{\delta}{\delta g_{\mu\nu}}\right)(\sqrt{g}\,\mathcal{L}), \tag{208}$$

where $\mathcal{L}_\xi$ is the Lie derivative along ξ, not to be confused with the Lagrangian (sorry!). Here the fields ϕ could be any sorts of fields (well, not spinors, for simplicity), so that if $\phi = A_\mu$ is a vector field, $(\mathcal{L}_\xi A)_\mu = \xi^\nu \partial_\nu A + \partial_\lambda \xi^\mu A^\lambda$. Now when we translate all the fields appearing in $\mathcal{L}$ along ξ, the Lagrangian changes by a total derivative, since we are just moving the Lagrangian infinitesimally along the flow:

$$\delta(\sqrt{g}\mathcal{L}) = \partial_\mu(\xi^\mu \sqrt{g}\mathcal{L}). \tag{209}$$

As a very simple check, consider a mass term for a scalar φ. Then

$$\mathcal{L}_\xi\varphi\frac{\delta}{\delta\varphi}(\sqrt{g}\varphi^2) = \sqrt{g}\xi^\mu\partial_\mu\varphi^2, \qquad (\mathcal{L}_\xi g)_{\mu\nu}\frac{\delta}{\delta g_{\mu\nu}}(\sqrt{g}\varphi^2) = \frac{1}{2}\sqrt{g}\varphi^2(g^{\mu\nu}\xi^\lambda\partial_\lambda g_{\mu\nu} + 2\partial\cdot\xi) \quad (210)$$

where we used that the Lie derivative is

$$(\mathcal{L}_\xi g)_{\mu\nu} = \xi^\alpha\partial_\alpha g_{\mu\nu} + \partial_\mu\xi^\alpha g_{\alpha\nu} + \partial_\nu\xi^\alpha g_{\mu\alpha}. \quad (211)$$

Now

$$\partial_\lambda\sqrt{g} = \frac{1}{2}\sqrt{g}g^{\mu\nu}\partial_\lambda g_{\mu\nu}, \quad (212)$$

and so

$$(\mathcal{L}_\xi g)_{\mu\nu}\frac{\delta}{\delta g_{\mu\nu}}(\sqrt{g}\varphi^2) = \varphi^2\partial_\mu(\sqrt{g}\xi^\mu). \quad (213)$$

Therefore the total variation is

$$\delta(\sqrt{g}\mathcal{L}) = \partial_\mu(\sqrt{g}\xi^\mu\varphi^2) \quad (214)$$

as expected.

A less trivial example is something which involves vectors and derivatives, like $\sqrt{g}A_\mu\partial^\mu\varphi$. The variations over the matter fields give

$$\left((\mathcal{L}_\xi A)_\mu\frac{\delta}{\delta A_\mu} + \mathcal{L}_\xi\varphi\frac{\delta}{\delta\varphi}\right)(\sqrt{g}A_\mu\partial^\mu\varphi) = \sqrt{g}\left(\partial^\mu\varphi(\xi^\nu\partial_\nu A_\mu + \partial_\mu\xi^\nu A_\nu) + A_\mu\partial^\mu(\xi^\nu\partial_\nu\varphi)\right), \quad (215)$$

while the variation over the metric produces, after some algebra,

$$(\mathcal{L}_\xi g)_{\mu\nu}\frac{\delta}{\delta g_{\mu\nu}}(\sqrt{g}A_\mu\partial^\mu\varphi) = (\xi^\lambda\partial_\lambda\sqrt{g}+\partial\cdot\xi)A_\mu\partial^\mu\varphi - A^\mu\partial^\nu\varphi\xi^\alpha\partial_\alpha g_{\mu\nu} - A_\lambda\partial_\alpha\xi^\lambda\partial^\alpha\varphi - A_\alpha\partial^\alpha\xi_\lambda\partial^\lambda\varphi. \quad (216)$$

Since $\partial_\alpha g_{\alpha\sigma} = -g_{\mu\nu}\partial_\alpha g^{\nu\lambda}g_{\lambda\sigma}$, the third term on the RHS is actually $+A_\mu\partial_\nu\varphi\xi^\alpha\partial_\alpha g^{\mu\nu}$. Adding up the contributions from the matter and the metric, we see that the last two terms on the LHS of the above equation cancel with two of the terms in the variation of the matter fields, again leaving us with

$$\delta(\sqrt{g}\mathcal{L}) = \partial_\mu(\sqrt{g}\xi^\mu A_\nu\partial^\nu\varphi). \quad (217)$$

Now we can see the general pattern that's at work: the variation over the metric produces derivatives of $\sqrt{g}$, ξ, and any $g_{\mu\nu}$s appearing in $\mathcal{L}$, plus some extra stuff coming from transforming the indices of the $g_{\mu\nu}$s. The variation over the matter fields produces derivatives of thins involving the matter fields, plus some extra stuff coming from the transformation of any vector indices that the matter fields have. The extra stuff from the matter fields and the extra stuff from the variation of the $g_{\mu\nu}$s cancel, since upper-index variations cancel lower-index ones. After the smoke clears, we are left with a total derivative.

Another way to see this is simply to write the Lagrangian as a d-form. If $L = \mathcal{L}\cdot\star 1$, where $\star 1$ is the volume form, then under the variation we have

$$\delta_\xi L = \mathcal{L}_\xi L = (i_\xi d + d i_\xi)L = d(i_\xi L), \quad (218)$$

where we have used Cartan's formula and that L is a top-dimensional form. Therefore the variation of the Lagrangian is indeed always a total derivative.

Since our diffeomorphism must vanish at ∂M, the upshot to the above discussion is that (no $\sqrt{g}$ in the integration measure on the LHS; it will be picked up from varying the action)

$$\int_M d^d x \left((\mathcal{L}_\xi g)_{\mu\nu} \frac{\delta}{\delta g_{\mu\nu}} + \mathcal{L}_\xi \phi \frac{\delta}{\delta \phi} \right) S[\phi, g_{\mu\nu}] = \int_{\partial M} d^{d-1} x^\mu \sqrt{g|_{\partial M}} \xi_\mu \mathcal{L}|_{\partial M} = 0. \tag{219}$$

Now by definition, (we are working in Euclidean signature; hence the sign)

$$\frac{2}{\sqrt{g}} \frac{\delta S}{\delta g_{\mu\nu}} = T^{\mu\nu}. \tag{220}$$

The conservation of currents associated to global symmetries holds on shell. So, choosing a configuration of matter fields which solves the EOM, we see that

$$\int_M d^d x \sqrt{g} \, (\mathcal{L}_\xi g)_{\mu\nu} T^{\mu\nu} = 0, \tag{221}$$

on shell.

To get a conservation law, we need to massage the Lie derivative slightly. First, add and subtract $\Gamma^\lambda_{\alpha\mu} g_{\lambda\nu} + \Gamma^\lambda_{\alpha\nu} g_{\mu\lambda}$ on the RHS of the expression (211) for the Lie derivative. The positive terms combine with the first term above to produce a covariant derivative of $g_{\mu\nu}$, which dies. The negative terms are the connection coefficients needed to turn the derivatives of ξ into covariant ones. Therefore

$$(\mathcal{L}_\xi g)_{\mu\nu} = \nabla_{(\mu} \xi_{\nu)}. \tag{222}$$

Since the stress tensor is symmetric, we then have

$$\int_M d^d x \sqrt{g} \, (\nabla_\mu \xi_\nu) T^{\mu\nu} = 0. \tag{223}$$

Now naively we would like to integrate by parts and conclude that T is covariantly conserved. This is correct, but nonzero work is required to demonstrate it. First, we need the product rule

$$\nabla_\mu (T^{\mu\nu} \xi_\nu) = (\nabla_\mu T^{\mu\nu}) \xi_\nu + T^{\mu\nu} \nabla_\mu \xi_\nu. \tag{224}$$

The first term on the RHS has two Christoffel symbol terms, which both have minus signs, while the second term on the RHS has one positive Christoffel symbol. Two of these cancel, leaving a single negative-sign Christoffel symbol, which just the right index structure to match the $\Gamma^\mu_{\mu\lambda} T^{\lambda\nu} \xi_\nu$ term on the LHS. Since this is straightforward algebra, and I won't write it out.

Now we need to argue that $\sqrt{g} \nabla_\mu (T^{\mu\nu} \xi_\nu)$ is a total derivative. This is only true because the covariant derivative is acting on a vector; if it was acting on a larger-rank tensor, it would not be a total derivative. Indeed, for any vector V^μ, we have

$$\begin{aligned} \frac{1}{\sqrt{g}} \partial_\mu (\sqrt{g} V^\mu) &= \partial_\mu V^\mu + \frac{1}{2} V^\mu g^{\alpha\beta} \partial_\mu g_{\alpha\beta} \\ &= \partial_\mu V^\mu + \frac{1}{2} V^\mu g^{\alpha\beta} (\Gamma^\lambda_{\mu\alpha} g_{\lambda\beta} + \Gamma^\lambda_{\mu\beta} g_{\alpha\lambda}) \\ &= \nabla_\mu V^\mu, \end{aligned} \tag{225}$$

where in the first line we used the usual way of differentiating the determinant by writing it as $e^{\frac{1}{2}\mathrm{Tr}\ln g}$, and in the next line we used the metric compatibility of the connection to substitute in some Christoffel symbols for the derivative of the metric. Therefore we indeed have

$$\sqrt{g}\nabla_\mu(T^{\mu\nu}\xi_\nu) = \partial_\mu(\sqrt{g}\,T^{\mu\nu}\xi_\nu). \tag{226}$$

Therefore, taking ξ^μ to vanish at ∂M, we conclude that

$$\int_M d^dx\,\sqrt{g}\,\xi_\nu\nabla_\mu T^{\mu\nu} = 0, \tag{227}$$

and since this must hold for any ξ_ν, we have

$$\nabla_\mu T^{\mu\nu} = 0, \tag{228}$$

as required.

Finally, we can take a look at the relationship between the stress tensor defined here and the "canonical" stress tensor $T_c^{\mu\nu}$ obtained from the Noether procedure. Under $\phi(x^\mu) \mapsto \phi(x^\mu + \xi^\mu)$, the action changes as

$$\delta S = \int_M d^dx\,\sqrt{g}\,\nabla_\mu\xi_\nu T_c^{\mu\nu}. \tag{229}$$

Since this variation is the one produced by the second term on the LHS of (219), we have, for any matter field configuration (on-shell or off),

$$\begin{aligned}\int_M d^dx\,\sqrt{g}\,\nabla_\mu\xi_\nu T_c^{\mu\nu} &= \int_M d^dx\,\sqrt{g}\,(\mathcal{L}_\xi g)_{\mu\nu}\frac{\delta}{\delta g_{\mu\nu}}S[\phi, g_{\mu\nu}] \\ &= \int_M d^dx\,\sqrt{g}\,\nabla_\mu\xi_\nu T^{\mu\nu}.\end{aligned} \tag{230}$$

Since this must hold for any ξ_ν, we may take the matter fields to be off-shell (so that $T^{\mu\nu}$ has a non-zero divergence), and conclude that

$$\nabla_\mu T_c^{\mu\nu} = \nabla_\mu T^{\mu\nu}, \tag{231}$$

which tells us that the two stress tensors agree up to something whose divergence is zero, namely the components of $d^\dagger B$, where B is a 3-form.

June 16 — Vacuum decay and unstable potentials

Today we'll be writing down some cool comments on unstable potentials given by Nima Arkani-Hamed at TASI 2019. There won't be much beyond just repeating what he said in lecture, but I wanted to have it written down for posterity's sake.

Solution:

We will be working in four dimensions throughout, for concreteness.

First, consider the potential $V(\varphi) = -m^2\varphi^2$, with $m^2 > 0$. This is clearly unstable, and dimensional analysis tells us that the decay time of the vacuum is

$$t_d \sim \frac{1}{m}. \tag{232}$$

Likewise, if $V(\varphi) = -\mu\varphi^3$, we have an unstable vacuum with a decay time of $t_d \sim 1/\mu$.

The tricky case is when $V(\varphi) = -\lambda\varphi^4$. There is no dimensionful parameter, so how are we supposed to determine the decay time? To think about this, let $\phi_R(x) = \phi_0\delta_{x\in R}$, where the indicator function $\delta_{x\in R}$ is 1 if $x \in R$ and 0 else.

$$\Psi[\phi_R(x)] = \exp\left(-\phi_0^2\right) \tag{233}$$

What is the energy of such a field configuration? We gain $g_n R^D \phi_0^n$ from the potential term, but we loose $(\phi_0/R)^2 R$ from the gradient term. Therefore the energy of the bubble is

$$V_R \sim -g_n R^D \phi_0^n + \phi_0^2 R... \tag{234}$$

This means that the critical size of the bubble $R_\star$ at which the bubble becomes energetically favorable to create is

$$R_\star \sim . \tag{235}$$

June 21 — The myriad appearances of the η invariant in discussions of anomalies

η invariants of various Dirac operators come up very often when discussing fermions and anomalies. Here we will try to sort out exactly where the η invariants of various Dirac operators appear in anomaly calculations.

Solution:

References