Notes on Abelian lattice gauge theories, quantum $\leftrightarrow$ classical mappings, and dualities

Ethan Lake

1 Notes on phase diagrams of Abelian lattice gauge theory

Today we're compiling a list of facts / intuitions about Abelian lattice gauge theories in 3 and 4 spacetime dimensions, both for quantum Hamiltonians in d spatial dimensions and their $(d+1)$-dimensional classical counterparts. We will be working at $T = 0$ throughout.[1] Most good references for background material are from back in the 70s (Fradkin + Shenker, Kogut + coworkers, etc.) and are all very classic + well-known, and so I haven't bothered to add citations. Thanks also to Bi Zhen and Ji Wenjie for discussions.

The quantum Hamiltonians we will be considering are variants of

$$H = -\frac{1}{2}\left(\frac{1}{g}\sum_{\square} B_{\square} + g\sum_l \mathcal{X}_l + \lambda \sum_l Z_i^\dagger \mathcal{Z}_l Z_{i+1} + \frac{1}{\lambda}\sum_i X_i\right) + h.c. \tag{1}$$

for the $\mathbb{Z}_N$ case and

$$H = -\frac{1}{2}\left(\frac{1}{g}\sum_{\square} \cos(B_{\square}) + g\sum_l E_l^2 + \lambda \sum_l e^{-i\phi_i} U_l e^{i\phi_{i+1}} + \frac{1}{\lambda}\sum_i \pi_i^2\right) + h.c. \tag{2}$$

for the $U(1)$ case. Here the notation is that $\mathcal{Z}_l$ and $U_l = e^{iA_l}$ are gauge link variables, $\mathcal{X}_l$ and E_l are the conjugate momentum variables to $\mathcal{Z}_l$ and A_l, and Z, X (ϕ, π) are the matter fields / momenta in the discrete (continuous) case.[2] The $B_\square$s are defined in the usual way: for each plaquette $\square$ we take a positively (wrt a reference orientation) oriented

[1] Thinking about $T > 0$ and the associated finite-temperature transitions in the phase diagrams of gauge theories is usually done with distinct methods (mapping to a spin model one dimension down via Polyakov loops) and hence will be relegated to a separated diary entry.

[2] Recall why we write the coefficients of the dynamical terms (g, λ^{-1}) as the inverses of the coefficients of the kinetic terms (g^{-1}, λ): this is a schematic way of indicating that in the Euclidean $D+0$ dimensional classical stat-mech model ("Lagrangian path integral"), the couplings appearing in the action are approximately the kinetic ones g^{-1}, λ. For situations where the fields are continuous (like $U(1)$), this relation (of the coefficient of the canonical momentum term getting inverted when passing to the stat mech model) is exact (up to the usual complications caused by the spacetime anisotropy of the corresponding classical model):

path around the plaquette, and add a factor of $\mathcal{Z}_l$ (or U_l) if the orientation of the path agrees with the orientation of the link (fixed by the orientation of the cubic lattice, which we will always be working on), or a factor of $\mathcal{Z}^\dagger/U^\dagger$ if the link disagrees with the orientation of the path. The generator of gauge transformations in the discrete case is $X_v \prod_{l\in\partial v} \mathcal{X}_l^{\pm_l}$, where $\mathcal{X}_l^{\pm_l}$ is $\mathcal{X}_l$ if l points into the vertex v and $\mathcal{X}_l^\dagger$ if l points out of v. For $U(1)$, it is the usual $\exp(i(\nabla\cdot E-\rho))$. In the $U(1)$ case, the matter fields we're working with are pure phase variables—their magnitudes are assumed to be frozen out by a large Higgs potential, regardless of what actual phase we are working in. The hope is that this can be done without unduly modifying the phase diagram.[3] If the full matter field is $\psi = |\psi|e^{i\phi}$, the above kinetic term is equal to (up to unimportant constants) $|\psi_i - U_l\psi_{i+1}|^2$ after the $|\psi|$s are frozen out.

Now numerical studies of gauge theories are done on D+0 dimensional classical stat mech models, rather than with quantum models that have continuous time. Therefore to compare with numerics, it's best to do the usual $\mathcal{QC}$ mapping on the above quantum Hamiltonian. This procedure is a bit ambiguous and different presentations of the Hamiltonian are more or less well-suited to doing the $\mathcal{QC}$ mapping. In fact the above Wilsonian form of the quantum Hamiltonian does *not* have a simple classical counterpart, and to do the $\mathcal{QC}$ mapping in a simple way it is necessary to work with the modified Villain form of the Hamiltonian. The classical partition function obtained from the Villain modification of the previous quantum H is the isotropic model

$$Z = \sum_{\substack{\{p_\Box\}\in\mathbb{Z}^{N_\Box},\\ \{p_l\}\in\mathbb{Z}^{N_l}}} \sum_{\substack{\{z_l\}\in\mathbb{Z}_N^{N_l}\\ \{\phi_v\}\in\mathbb{Z}_N^{N_v}}} \exp\left(-\frac{1}{2g}\sum_\Box\left(\frac{2\pi}{N}\mathcal{H}_\Box - 2\pi p_\Box\right)^2 - \frac{\lambda}{2}\sum_l\left(\frac{2\pi}{N}((d\phi)_l - z_l) - 2\pi p_l\right)^2\right) \tag{4}$$

this is basically due to the relation that

$$\int dx'\, e^{-(x-x')^2/2K} f(x') = e^{-K\nabla^2/2} f(x), \tag{3}$$

which follows from writing f in fourier components and doing the quadratic integral (basically what's happening when going from Hamiltonians to Lagrangians—integrating out the momentum, which we assume to appear quadratically in the action). Note here that the coefficient of the canonical momentum K becomes K^{-1} when we write things in terms of the integral over "time" (the x coordinate in the above). Hence e.g. for Maxwell theory, the Hamiltonian $g^{-1}B^2 + gE^2$ is equivalent to a stat-mech model in one dimension up with Hamiltonian $g^{-1}F^2$.

Now when the fields aren't continuous (or even if they are but we aren't in the continuum limit), this nice relation between the coefficients of the stat mech model and the quantum Hamiltonian doesn't hold—recall from the example of the transfer matrix in the Ising model that the coefficients in the classical model are written in terms of hyperbolic trig functions and such. So while in general it is incorrect to find the corresponding stat mech model by inverting the coefficient of the dynamic term, this rule is schematically correct, and at any rate when discussing stat mech models we will always assume that after doing the quantum → classical mapping, the model can be made homogeneous in spacetime and the precise coupling constants adjusted slightly without taking us out of the universality class in question.

[3]Since the gauge fields only couple to the phase degree of freedom, this seems like a reasonable expectation.

Figure 1: From left: an open string worldsheet, a schematic representation of the partition function, and a schematic representation of the ground-state wavefunction.

in the $\mathbb{Z}_N$ case and

$$Z = \sum_{\substack{\{p_\square\}\in\mathbb{Z}^{N_\square},\\ \{p_l\}\in\mathbb{Z}^{N_l}}} \int \prod_l \mathcal{D}z_l \exp\left(-\frac{1}{2g}\sum_\square (\mathcal{H}_\square - 2\pi p_\square)^2 - \frac{\lambda}{2}\sum_l ((d\phi)_l - z_l - 2\pi p_l)^2\right) \tag{5}$$

in the $U(1)$ case. Here the ps are integers used to write the Villain approximation to the cosines appearing the Wilsonian form of the action, and the d is a lattice differential. The $\mathcal{H}_\square$s are the gauge holonomies around $\square$; they are valued in $\mathbb{Z}_N$ or $[0, 2\pi)$ (and not $\mathbb{R}$!), as appropriate. While the Villain form looks complicated it is actually much simpler than the usual Wilsonian form (which is hoped to be in the same universality class as the above) due to the fact that all the fields appear quadratically. Note that the coefficients g, λ here are *not* the same as the ones in the quantum Hamiltonian (they don't even have the same dimension as the ones in the quantum Hamiltonian, nor is the relation between the two sets of coefficients very simple in general). That said the g, λ here control the same physics as the g, λ in the quantum Hamiltonian, and we will as such continue with this abuse of notation.

In what follows we will mostly discuss the quantum (Wilsonian) Hamiltonian, but when comparing with numerics will refer to the above classical action. Again, the hope is that the two approximations made (changing from the Wilsonian quantum Hamiltonian to a Villain approximation, and then approximating the resulting classical model by one that's homogeneous in spacetime) don't qualitatively affect the structure of the phase diagrams.

We will make some general comments on the 2+1D and 3+1D cases first, and then do a more systematic discussion of the phase diagrams. Until discussing the phase diagrams, all matter will be assumed to have unit charge.

To get an intuitive picture on the various phases, it's helpful to look at the D-dimensional stat mech model (alias the path-integral representation) in order to get some geometric intuition. For simplicity, consider the $\mathbb{Z}_2$ case, and (for the rest of this section only) consider the Wilsonian version of the classical model, viz.

$$S = -\frac{1}{2g}\sum_\square B_\square - \frac{\lambda}{2}\sum_{\langle ij\rangle} Z_i \mathcal{Z}_{\langle ij\rangle} Z_j, \tag{6}$$

where the $Z_i, \mathcal{Z}_l$ are now just numbers. The specification to $\mathbb{Z}_2$ makes writing the partition function easy since $e^{\alpha s} = \cosh\alpha + s\sinh\alpha$ for $s \in \mathbb{Z}_2$. Hence

$$Z \propto \sum_{\{\mathcal{Z}_l, Z_i\}} \prod_\square (1 + \tanh(g^{-1})B_\square) \prod_{\langle ij\rangle} (1 + \tanh(\lambda) Z_i \mathcal{Z}_{\langle ij\rangle} Z_j). \tag{7}$$

The terms which survive this sum are then terms that are formed by surfaces built from $B_\square$s, which are either closed or which terminate on $Z\mathcal{Z}Z$ lines. The tendency for these surfaces to proliferate, as well as whether or not they tend to be open or closed, depends on the surface tension $\sim [\tanh(g^{-1})]^{-1}-1$ and the line tension $\sim [\tanh(\lambda)]^{-1}-1$. The ground-state wavefunction $|\Psi\rangle$ is found by cutting open the path integral along a spatial slice; hence the ground state is made from a combination of open and closed strings. The case of $D=3$ is shown in the top row of Figure 1.

2+1D

Pure gauge theory

First we consider the $\mathbb{Z}_N$ case with $\lambda = 0$. First consider the deconfined regime of the quantum model. At $g=0$ we need $B_\square$ to act as $\mathbf{1}$ on every plaquette. The ground state here is a gas of closed strings of all sizes; this follows from cutting open the partition function (7) and noting that at $g=0$ the surface tension vanishes while the line tension is infinite.

We can construct the $g=0$ ground state $|\Omega\rangle$ explicitly either by applying all possible t'Hooft lines[4] to the $\otimes|\uparrow\rangle$ product state, or by applying all possible Wilson lines to the $\otimes|+\rangle$ product state:

$$|\Omega\rangle = \sqrt{N}^{N_\square}\prod_\square \Pi_{B_\square}|\otimes +\rangle = \sqrt{N}^{N_\square}\prod_v \Pi_{A_v}|\otimes\uparrow\rangle \tag{10}$$

where the projector $\Pi_\mathcal{O}$ for an operator $\mathcal{O}$ with eigenvalues in the Nth roots of unity is defined as the projector onto the $+1$ eigenspace:

$$\Pi_\mathcal{O} = \frac{1}{N}\sum_{k\in\mathbb{Z}_N}\mathcal{O}^k. \tag{11}$$

Here also A_v is the usual $\prod_{l\in\partial v}\mathcal{X}_l$ vertex operator (the generator of gauge transformations in the pure gauge theory). Just to check the normalization and that the two ways of writing the state are indeed the same, we check

$$\langle\Omega|\Omega\rangle = N^{N_\square}\frac{1}{N^{2N_\square}}\langle\otimes\uparrow|\prod_{v,\square}\sum_{k,l\in\mathbb{Z}_N}A_v^kB_\square^l|\otimes +\rangle. \tag{12}$$

Since $B_\square$ is gauge invariant the A_vs and $B_\square$s commute; hence

$$\langle\Omega|\Omega\rangle = N^{N_\square}\langle\otimes\uparrow|\otimes +\rangle = N^{N_\square - N_l/2} = 1, \tag{13}$$

[4] In two spatial dimensions, t' Hooft lines are defined by

$$T_\gamma = \prod_{l\in\gamma}\mathcal{X}_l \tag{8}$$

for γ any path in $C_1(\Lambda^\vee;\mathbb{Z}_N)$, where $\Lambda^\vee$ is the dual of the spatial lattice Λ. To talk about dynamical questions, which is all we're really going to be caring about, we can focus on spatial manifolds with trivial topology. In this case we can always write

$$T_\gamma = \prod_{v\in\bar\gamma}A_v, \tag{9}$$

where $\bar\gamma$ is the collection of all vertices enclosed by γ.

where N_l is the number of links.

First, note that W_γ has unit expectation value for all choices of γ. A perimeter law only sets in for W_γ when we add some line tension. When we turn on a small $g > 0$, we add in processes which measure electric flux, i.e. which create magnetic flux. Since the string of the Wilson line detects magnetic flux, dynamically created magnetic flux leads to nonzero line tension (said another way, the Wilson line creates an electric flux tube, which is now explicitly tensionful).

Near the Ω state, we determine the vev of the Wilson line by doing perturbation theory. The eigenstates of the small g Hamiltonian which enter the expansion at nth order in perturbation theory are those which contain n pairs of fluxes on adjacent plaquettes, which can be written as $\prod_l \mathcal{X}_l|\Omega\rangle$ for some set of links l. Therefore if $|\Omega_g\rangle$ is the perturbed ground state,

$$\langle W_\gamma\rangle = \frac{\langle\Omega_g|W_\gamma|\Omega_g\rangle}{\langle\Omega_g|\Omega_g\rangle} = 1 + \frac{1}{4}\sum_{l,l'} g^2\langle\Omega|\mathcal{X}_l^\dagger W_\gamma \mathcal{X}_{l'}|\Omega\rangle + h.c. + \dots, \tag{14}$$

where $\dots$ come from fluxes separated by multiple lattice spacings and configurations with more than two fluxes. If the length of the Wilson line $|\gamma|$ is large enough, configurations with flux pairs separated by multiple lattice spacings can be ignored, and the flux pairs (flipped links) can be treated as a dilute gas of bosons. This means that the coefficient of the nth order term in perturbation theory goes as $(g^2/4)^n/n!$, so that the above series re-exponentiates and gives

$$\ln\langle W_\gamma\rangle = -\frac{g^2}{4}|\gamma|, \tag{15}$$

up to corrections coming from approximating the flux pairs as dilute, which will be exponentially small in $|\gamma|$ (apparently it can be proven that this series expansion has a finite radius of convergence in g). This is the expected perimeter law.

The t Hooft lines T_γ also have unit expectation value in $|\Omega\rangle$. In fact they have exactly unit expectation value in the pure gauge theory even when the line tension is turned on—this is because the string of the t Hooft line is moved by applying gauge transformations A_v; hence the location of the string is unphysical and it cannot possibly have any tension. Therefore in the pure gauge theory, the t' Hooft lines (by "line" we really always mean "loop"; open t' Hooft lines have energy density at their ends) are essentially trivial and hence there is no meaningful magnetic operator. This is in keeping with the fact that for lattice $\mathbb{Z}_N$ gauge theories, the magnetic 1-form symmetry is always explicitly broken by the lattice (it only emerges when we work strictly at $g = 0$, where excitations with nonzero field strength are projected away), and so the t' Hooft loops can never be used as a diagnostic of what phase we're in.

Now let's go to the confining regime. The $g = \infty$ groundstate $|\Gamma\rangle$ is just a $\otimes$ of $+1$ eigenstates of $\mathcal{X}_l$ on every link. That the Wilson line obeys an area law is seen just from the usual high temperature expansion: expanding the $g^{-1}B_\square$ exponential[5] we see that the first

[5] I guess since the two gauge terms in H don't commute, it's probably better to first go to the classical model one dimension up and do the expansion of the exponential there.

term that survives has a product of $B_\Box$ filling the loop—therefore

$$\frac{\langle \Gamma_g | W_\gamma | \Gamma_g \rangle}{\langle \Gamma_g | \Gamma_g \rangle} \sim e^{A(\gamma) \ln g^{-1}}, \tag{16}$$

giving the expected area law.

The $U(1)$ case is different, since in this case we have monopoles, which prohibit a deconfined / topological phase. The intuition here is that in the 3+0D stat mech model (path integral), the operators which disorder the Wilson loops are monopoles, which are 3+0D are point-like instantons. It is therefore always favorable (in the free-energy sense) to have a macroscopic number of monopoles, which disorder the system and prevent the Wilson lines from getting a perimeter-law expectation value. This is basically the same as the free energy argument for why domain walls always disorder models with discrete symmetry in 1+0D—if the disordering operator is point-like, free energy considerations always favor proliferating it. If the disorder operator acting in a putative ordered phase costs energy E, then the free energy for configurations with a single disorder operator in a volume L goes as $F \sim E - \alpha \ln L$ for some constant α. When $L \to \infty$ we therefore always get a disordered phase at long distances (although we have to go to distances exponentially large in E [i.e. flowing for an RG time $\propto E$] in order to see the disordering). This argument also applies to $\mathbb{Z}_N$ gauge theory in 2+0D: in this case the disorder operators are just $\mathcal{X}_l$, which being local mean that this theory is always confined.

The Wilson line changes from a marginally confining QLRO form of $\langle W_\gamma \rangle \sim e^{-g^2 \ln |\gamma|}$ in the free-field approximation at small g to the regular area law at large g. Given that the Wilson line scales as a power of $|\gamma|$ in the free-field approximation, this seems to suggest that there's a KT transition between the power law and the exponential decay at large g. However in fact unlike the XY model, at any $g > 0$ the theory is massive, with $\star F$ having a mass $\propto e^{-1/g^2}$ (the action of the instanton)—this means that the KT transition occurs at $g = 0$ (although perturbatively in g the mass is always zero). It is useful to contrast this with the case of the XY model—in both cases, the action can be dualized and written as

$$S = \int d^d x \left(\frac{R^2}{4\pi} |d\phi|^2 + y \cos \phi \right), \tag{17}$$

where y depends on the action of the instanton and $R^2 \sim 1/g^2$ for the gauge theory. The difference is that in $D = 2$ the relevance of the cosine depends on the value of R^2, while for $D = 3$ the cosine is always relevant. Hence $U(1)$ gauge theory is massive at all nonzero g, while the XY model has a nonzero regime of masslessness.

With matter

Now turn on the coupling to matter, and again focus on the $\mathbb{Z}_N$ case. A useful tool for understanding the phase diagram is EM duality. This works in the quantum model via (as usual $\vee$ denotes cells on the dual lattice)

$$B_\Box \leftrightarrow X_{\Box^\vee} \qquad \mathcal{X}_l \leftrightarrow Z^\dagger_{(\partial l^\vee)_0} \mathcal{Z}_{l^\vee} Z_{(\partial l^\vee)_1}. \tag{18}$$

That is, the kinetic term for the gauge fields goes to the dynamic (momentum) term for the matter fields, and vice versa. Therefore in 2+1D the discrete lattice theory has a natural notion of EM duality sending $g \leftrightarrow \lambda$, implying a symmetry of the phase diagram. Under the duality, Wilson lines map to electric flux operators:[6]

$$\prod_{l\in\gamma} \mathcal{Z}_l \leftrightarrow \prod_{l\in\gamma^\vee} \mathcal{X}_l, \tag{19}$$

so that electric and magnetic fluxes are exchanged, as usual. Note that the RHS can be re-written as $\prod_{v\in\bar{\gamma}} X_v$ with $\bar{\gamma}$ the region enclosed by γ. Hence the Wilson line (the order parameter for the gauge theory) maps to the disorder operator on $\bar{\gamma}$ in the spin system; this is also what we expect from EM duality.

Since electric matter can screen the Wilson line, the latter follows a perimeter law at any nonzero λ. It is obviously still a perimeter law for small g, λ, while for large g we can do the high temperature expansion and expand the exponential in both g^{-1} and λ: there is always a term coming from $(Z_i^\dagger \mathcal{Z}_{\langle ij\rangle} Z_j + h.c.)^{|\gamma|}$ that contributes to $\langle W_\gamma\rangle$, and so in the limit of $|\gamma| \to \infty$, the leading behavior of the Wilson line will be $\langle W_\gamma\rangle \sim e^{-\ln(\lambda^{-1})|\gamma|}$. Nevertheless there is still an open Wilson line that can be used to distinguish the two phases at $\lambda > 0$; this is discussed in a subsequent section.

Just as dynamical magnetic flux is responsible for giving tension to Wilson lines, dynamical electric matter gives tension to t' Hooft lines; this is possible since since now the operator which moves the t' Hooft lines, viz. A_v, no longer generates gauge transformations. The fact that electric matter makes the t' Hooft lines tensionful is another way of explaining why charged matter fights confinement. Anyway, we check the tension at small λ by computing $\langle T_\gamma\rangle$ in an expansion in powers of λ. The ground state has all the matter fields in the $+1$ eigenstate of all the X_v operators, and the perturbations to this state come from states where pairs of electric charges are excited across a given link. This gives a finite line tension $\propto \lambda^2$ by the same calculation as (14).

Since we always have dynamic magnetic flux (created by the $g\sum_l \mathcal{X}_l$ term), the t' Hooft line can always be screened, and hence will always follow a perimeter law, even in the deconfined regime when $g \to 0$.[7] This again follows from a high temperature expansion, which tells us that $\langle T_\gamma\rangle \sim e^{-|\gamma| g}$ in the deconfined phase.

For the $U(1)$ case, adding (unit charge) matter doesn't really do anything. As we said, the fact that the pure gauge theory always confines is due to the presence of a photon mass induced by instantons. The mass term is relevant, and so in order for matter to do anything, its presence would need to change the scaling dimension of the gauge field considerably. This may be possible at intermediate coupling (or in large N), but I think the expectation is that for a single species of matter, the spectrum is still confined and massive at all λ.

[6] For economy of notation we aren't bothering to keep track of orientations of links and possible †s that need to be added; if the reader is bothered by this just pretend we are in the $N = 2$ case.

[7] Now in 3+1D, a phase in which both W and T are P-law implies the existence of gapless degrees of freedom—the correlation of linked W and T operators differs from that of unlinked ones, even though both operators are perimeter law, and this is only possible if there is a long-ranged force that mediates communication between the two operator insertions. In 2+1D this argument doesn't apply to the deconfined phase where both W and T are P-law, since there are no configurations of W and T loops that link / have a nontrivial signed intersection number (we are on a space of trivial topology).

3+1D

Pure gauge theory

In 3+1D the $\mathbb{Z}_N$ pure gauge theory has an EM duality, since the dual of a link (electric field) is a plaquette (magnetic field). This duality does $B_\square \leftrightarrow \mathcal{X}_l$, and hence does $g \leftrightarrow g^{-1}$.

This duality means that e.g. if there is only a single phase transition between the confined and topological phases, it must occur at $g = 1$. Furthermore if there are two transitions, they must be related by inversion about $g = 1$. The former possibility is observed to occur for $N = 2, 3, 4$, where 4+0D lattice calculations find a first-order transition at the coupling corresponding to the choice $g = 1$ in the quantum model.[8]

The duality is simple to see in the Hamiltonian model, but most of our knowledge of the phase diagrams comes from numerics, which are done on 4+0D lattice models. As shown in another diary entry, the classical pure gauge theory with action (4) is dual to a copy of itself with gauge coupling

$$\frac{1}{\widetilde{g}} = \frac{gN^2}{4\pi^2}. \tag{20}$$

Therefore if there is only one transition, in the classical gauge theory it must occur at

$$g_* = \frac{2\pi}{N}, \tag{21}$$

so that larger N theories have "more" confinement. In particular, the self-dual point moves to weaker coupling as N gets large.

The behavior of the Wilson lines deep in both phases is calculated using the same expansions as in the 2+1D case. The t' Hooft operator is now defined as

$$T_{\Sigma^\vee} = \prod_{l \pitchfork \Sigma^\vee} \mathcal{X}_l. \tag{22}$$

In Euclidean spacetime, this corresponds to changing the signs of all temporally-oriented plaquette terms whose "bottom" (with respect to the time direction) link is included in the above product. More generally we can also consider t Hooft operators with support on manifolds $\Sigma^\vee$ not contained within spatial slices; in Euclidean spacetime these insertions correspond to changing the signs of the plaquette terms for all plaquettes dual to $\Sigma^\vee$.

Since we are on a lattice where there is always magnetic matter, the dual surface $\Sigma^\vee$ is always allowed to have a nonempty boundary.[9] If $\partial\Sigma^\vee \neq \emptyset$, $T_{\Sigma^\vee}$ creates a loop of magnetic flux along $\partial\Sigma^\vee$. Just as in 2+1D, in the pure gauge theory the interior of the t Hooft operator

[8] This is actually rather surprising since the quantum-to-classical mapping only rigorously works when g is small.

[9] Note that we do not want to define the t Hooft operator as

$$T_{\Sigma^\vee} = \prod_{l \in \partial\Sigma^\vee} \mathcal{X}_l, \tag{23}$$

(where $\partial\Sigma^\vee$ contains the boundary links of those links which intersect $\Sigma^\vee$ transversely) — the t Hooft operator in 3+1D $\mathbb{Z}_N$ gauge theory is a surface operator not a line operator, since it needs to be able to link with Wilson lines in four-dimensional spacetime.

can be moved by gauge transformations. This means that if $\Sigma^\vee \in B_2(\Lambda^\vee;\mathbb{Z})$ (here $\Lambda^\vee$ is the dual lattice), then $T_{\Sigma^\vee}$ is gauge-equivalent to $\mathbf{1}$. If $\Sigma^\vee$ is nontrivial in $H_2(\Lambda^\vee;\mathbb{Z})$ it acts as a nontrivial operator, although is completely tensionless regardless of what phase we are in. We might have thought that in 3+1D we would be asking whether the t' Hooft operators satisfied volume law or area law, but in fact the interesting t Hooft operators are actually the open ones where $\partial\Sigma^\vee \neq \emptyset$, which will satisfy either area law or perimeter law.

In the deconfined phase, the leading contribution to $\langle T_{\Sigma^\vee}\rangle$ with $\partial\Sigma^\vee \neq \emptyset$ is obtained by expanding $e^{-g\sum_l X_l}$ until one gets a configuration with no magnetic flux. This happens at order $g^{|\Sigma^\vee|}$, and so the open t Hooft operators obey an area law. In the confined phase, the same sort of calculation as before gives a perimeter law with coefficient proportional to $1/g$.

In the $U(1)$ case, there is no duality to help us—the theory has dynamic magnetic matter built in (by virtue of the fact that we're on a lattice) but lacks electric matter, and as such cannot possibly be self-dual[10] (this is indicated by the fact that the classical gauge theory duality just discussed becomes $g \overset{?!}{\leftrightarrow} 0$ in the $N\to\infty$ limit).

Anyway, unlike in 2+1D, magnetic monopoles fail to disorder the system at all g. This is basically just because the mangetic monopoles now form strings in spacetime, and so the free energy arguments about their proliferation are modified. Indeed, consider a t' Hooft line T of length L, with L macroscopically large, i.e. of the same order as the size of the system. We want to ask when such lines proliferate and disorder the system; this is done by computing their free energy in the 4+0D stat mech model. Since there is always magnetic matter around to screen them, their average energy will behave as a perimeter law, with $E_T = CL$ for some $C(\lambda, g, \dots)$. The configurational entropy that's relevant here is the log of the number N_L of loops of length $\sim L$ in a box of size L^4. This is known to scale exponentially in L, i.e. $N_L \sim (C')^L$ for some constant C'. The relevant free energy is then

$$F_T \sim CL - \alpha \ln(C')^L \sim L(C - \alpha\ln(C')) \tag{24}$$

for some α which again is a function of g,λ. Whether or not the t' Hooft lines proliferate is then a question of what the functions C, C', α are. Therefore a Coulomb phase is possible unlike in the 2+1D case where the sign of the free energy is always negative at large L.[11]

So, we know that $U(1)$ lattice gauge theory in 4+0D has a Coulomb phase at small g and a confined phase at large g. The simplest possibility is that there is a single transition between the Coulomb phase at $g\to 0$ and the confined phase, which is indeed borne out by numerics. Therefore if the $N\to\infty$ limit of the $\mathbb{Z}_N$ theories is to mimick the $U(1)$ theory, there cannot always be just a single transition in the $\mathbb{Z}_N$ theories, since the phase diagrams of the $\mathbb{Z}_{N\to\infty}$ and $U(1)$ cases wouldn't match up. As N increases there must then be a third Coulomb phase which springs up and eventually subsumes the topological phase. If we assume that there are just these three phases, then the phase boundaries are related by (21):

[10] Since we're on a lattice we always have to sum over configurations of U_ls with $\prod_{\square\in\partial\text{⧉}} B_\square \neq 1$ around certain cubes. Such cubes would map to dual vertices for which the gauge transformation generator A_v doesn't act as $\mathbf{1}$, and so in order to get the duality to work we'd need electric matter as well.

[11] This is the same argument used to show that $\mathbb{Z}_N$ theories in 2+1D can have a deconfined phase at finite g,λ, since in that case the disorder operators are also one-dimensional (in that case they are open lines; closed t' Hooft loops commute with all Wilson lines and hence do not disorder the system like they do in 3+1D).

	$D = 2+1$			$D = 3+1$		
	$t=0$	$t=1$	$t=2$	$t=0$	$t=1$	$t=2$
	$t=0$	$t=1$	$t=2$			
$dir = 1$						

Table 1: Caption

the confined-Coulomb phase boundary moves to the critical coupling for the $U(1)$ transition, and the Coulomb-topological boundary moves to $g = 0$ as $N \to \infty$. The value N_* of N for which the Coulomb phase appears can be bounded by looking for when the self-dual point moves past the critical point of the $U(1)$ theory; it turns out that $N_* = 5$.[12]

With matter

Again, the computations of the scaling properties of the Wilson and t' Hooft lines is essentially the same as in the 2+1D case, so we won't elaborate much. With matter turned on the duality that we used in the $\mathbb{Z}_N$ case isn't so helpful, since it maps the theory onto something with 2-form gauge fields coupled to 1-form matter. The $U(1)$ theory now admits EM duality, but this is not so helpful since the treatment of magnetic matter in this formulation is a bit awkward.

When electric matter is turned on, the scaling of the t Hooft operators changes. The surfaces of t Hooft operators can no longer be moved using gauge transformations. Since the theory includes magnetic matter, the t Hooft operators on closed surfaces will never satisfy a volume law. However, t Hooft operators on open surfaces will always satisfy an area law, and never a perimeter law. This is because in the confined regime we still have electric charges which can be nucleated at short distances and detect the surface of the t Hooft operator. Indeed, at small $1/g$ and λ, we get

$$\langle T_{\Sigma^\vee} \rangle \sim 1 + \left\langle \left(\frac{1}{g} \sum_{\square} B_\square + \lambda \sum_l Z_i^\dagger \mathcal{Z}_l Z_{i+1} \right) T_{\Sigma^\vee} \left(\frac{1}{g} \sum_{\square'} B_{\square'} + \lambda \sum_{l'} Z_i^\dagger \mathcal{Z}_{l'} Z_{i+1} \right) \right\rangle + \cdots , \tag{25}$$

where the $\cdots$ are higher order terms. Even if λ is much smaller than $1/g$, the number of $Z^\dagger \mathcal{Z} Z$ terms which don't commute with $T_{\Sigma^\vee}$ goes as $|\Sigma^\vee|$, and the dominant contribution to the expectation value as $|\partial\Sigma^\vee|/|\Sigma^\vee| \to 0$ is an area law.

Of course, the area law only holds in the How does this then agree with the fact that in the Coulomb phase occurring at large enough N, the t Hooft operator in the $U(1)$ formulation obeys a perimeter law?

A summary of the behaviors of the various types of operators is shown below:

[12] This argument guarantees that we are not in the pathological situation where the behavior of $\mathbb{Z}_{N\to\infty}$ is different from that of $U(1)$.

Phase diagrams

The phase diagrams for the $\mathbb{Z}_N$ case are shown in Figure 2. We label the deconfined region as "topological" instead of deconfined to distinguish it from the Coulomb phase, which also has deconfined charges but is massless, unlike the topological phase.[13]

First look at the top left phase diagram. Here the symmetry of the phase diagram is determined by the duality $\lambda \leftrightarrow g$.[14] The second order lines and the first order part past the tricritical point are known from numerics, and their existence and exponents can be established near the edges by using the known properties of the $\mathbb{Z}_N$ matter model at $g = \infty$. The curvature of the lines is determined in perturbation theory: integrating out the matter at small λ adds gauge plaquette terms to the action, hence producing $1/g_{eff} = 1/g + C\lambda^4 + \dots$ with $C > 0$; the transition is therefore shifted to larger g. This makes sense—matter fields fight confinement. Likewise, the effect of gauge fluctuations in the spin model is to promote disorder, since they reduce the energy cost of having non-parallel spins due to the fact that they make the sign of the spin-spin coupling fluctuate slightly. Hence they increase the critical value of λ, in accordance with the duality.

The top right shows the phase diagram when the matter has $q \in \mathbb{Z}_N$, with $\gcd(q, N) > 1$. The top part of the phase diagram is the bottom part of the $\mathbb{Z}_q$ phase diagram, since there we have pure $\mathbb{Z}_q$ gauge theory. Since $q < N$, the Higgs transition meets the $\lambda = \infty$ axis at a larger value of g. Not sure what happens to the formerly first-order part after the tricritical point.

The bottom left is in 3+1D, with unit charge matter, and for $N \geq 5$ where a Coulomb phase is present. As $N \to \infty$ this Coulomb phase subsumes the topological phase; for $N < 5$ only the topological and confined phases are present. The two vertical weakly first-order transitions are related by duality and are likely first-order due to the fact that the transition from Coulomb to Higgs near $\lambda = 0$ involves condensing unit strength monopoles, which by HLM we expect to be weakly first order (and hence by duality the transition into the topological phase is weakly first-order as well). Similarly the upper green line involves the condensation of electric charges a la the usual Higgs mechanism, which is weakly first order for the same reason. The blue line is a guess: I think that the absence of a massless photon in the topological phase means the HLM mechanism won't be effective, allowing the second order transition in the pure matter theory to survive the presence of the gauge field.

The bottom right is for $q > 1$ matter. Nothing really new here: we just take the bottom part of the $\mathbb{Z}_q$ phase diagram and paste it on top.

The case with gauge group $U(1)$ is shown in Figure 3. The top left is fundamental matter in 2+1D, which is boring as Polyakov taught us. When $q > 1$ we know we need to get a $\mathbb{Z}_q$ phase diagram on the top of the diagram; presumably the situation is something like the one drawn.

The bottom left shows 3+1D. Here the HLM mechanism (either for electric charges or monopoles, as the case may be) implies that both transitions near the edges are weakly first

[13] Looking at correlators of $B_\square$ operators shows that the $\mathbb{Z}_N$ theories are always massive in the large g and small g limits.

[14] This is how the couplings get mapped in the quantum Hamiltonian, but *not* how they are mapped in the classical lattice model, where the transformation involves a factor of N (see the diary entry on $\mathcal{QC}$ mappings).

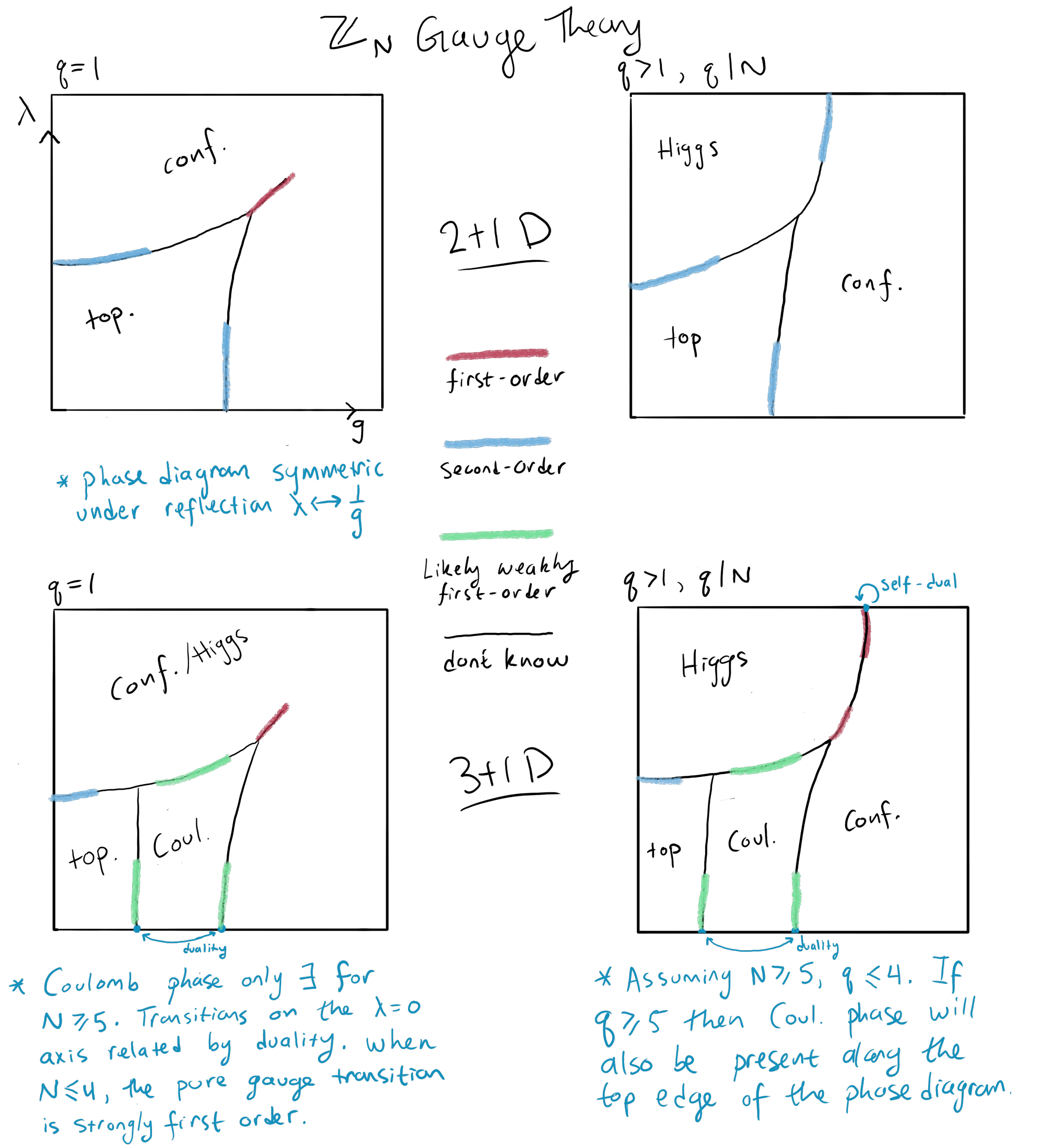


Figure 2: See text for comments. Top left should say reflection $\lambda \leftrightarrow g$.

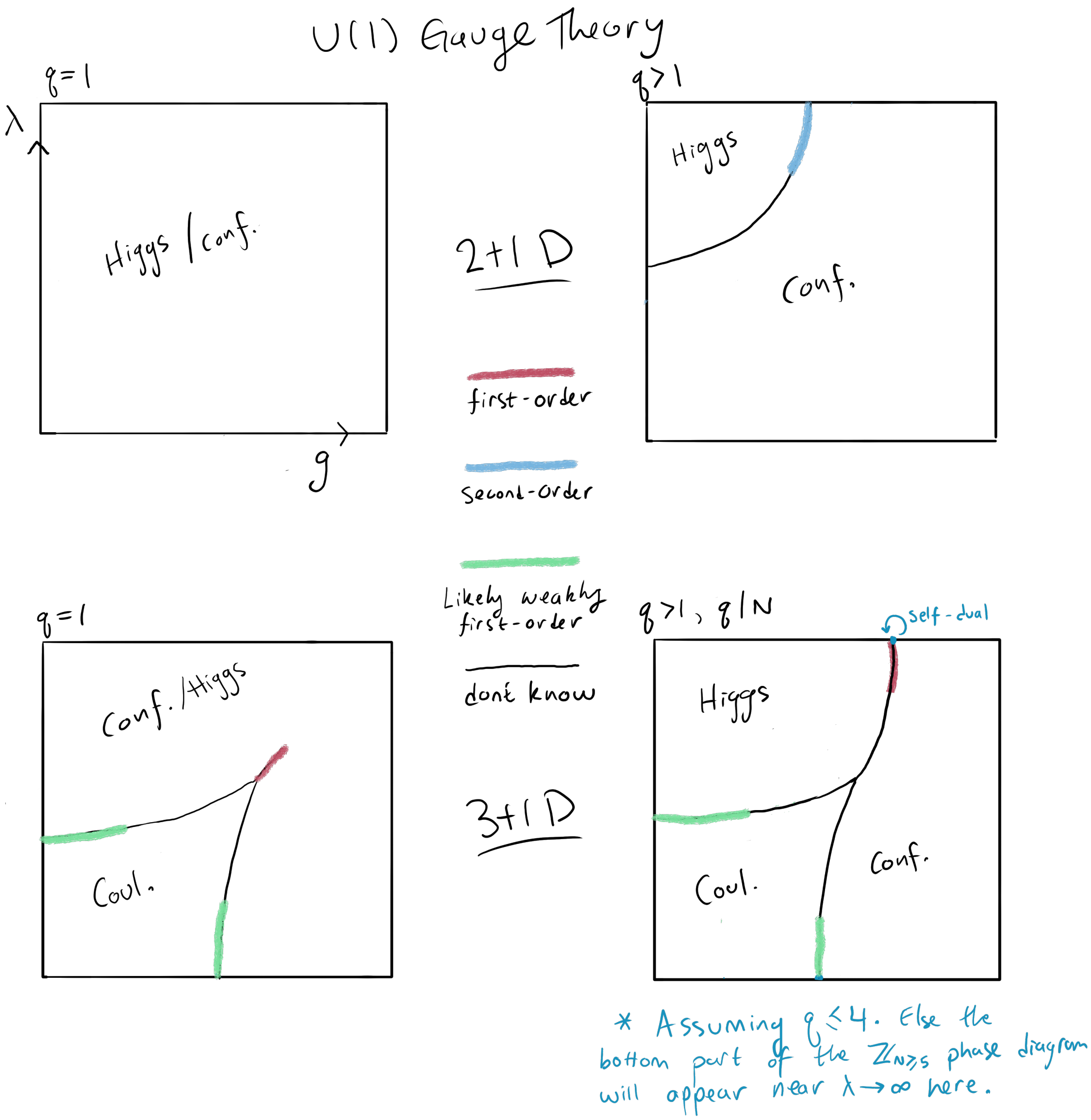


Figure 3: See text for comments.

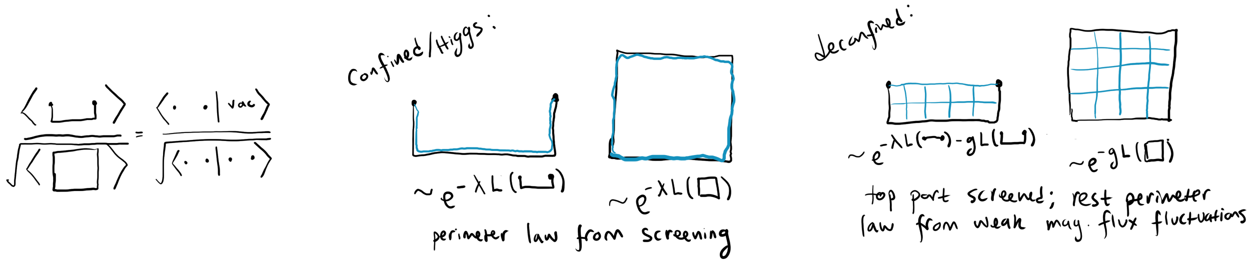


Figure 4: From right to left: a picture for the scaling of the cup operator in the deconfined phase, the same thing in the confined / Higgs phase, and a schematic rewriting of the ratio of Wilson lines as an inner product between a state with two charges and the vacuum.

order. For $q > 1$ as in the bottom right diagram, the story is as with $q = 1$ except with a copy of the $\mathbb{Z}_q$ gauge theory phase diagram along the top part.

Distinguishing the phases with open Wilson lines and the better "definition" of confinement

Given that when $\lambda \neq 0$ both W and T are always P-law, how do we distinguish the confined / Higgs phase from the deconfined one?[15] One answer was provided in the very cool paper [1]; this paper was sufficiently enlightening for me that I figured it would be worth taking a subsection to summarize it.

The point is that one should use a different order parameter, which measures where the line tension is coming from: is it coming from fluctuating magnetic fields around the path of the Wilson line (deconfinement), or is it coming from the energy needed to create the electric charges which screen the Wilson line (confinement)? There are many such operators which can tell the difference between these two types of line tension; the simplest one is the expectation value of the "cup operator" (my dumb name for it), defined by

$$C(L) = \frac{\langle Z_i^\dagger \left(\prod_{l\in\sqcup} \mathcal{Z}_l\right) Z_{i+L}\rangle}{\sqrt{\langle \prod_{l\in\square} \mathcal{Z}_l\rangle}}. \tag{26}$$

Here $\sqcup$ is a path in the shape of a $\sqcup$ with ends at the lattice sites $i, i+L$ (the addition is along some arbitrary direction), and is such that the two sides with ends are of length $L/2$, with the bottom segment having length L. The path $\square$ is a square with all sides of length L.

The meaning of this order parameter is in the left panel of figure 4. If we orient the vertical half-links of the cup in the time direction, then thinking in terms of a cut-and-glue approach to the partition function, we see that $\langle\sqcup\rangle$ represents the overlap between the vacuum (what you get when you cut open the $i\mathbb{R}$ time path integral) and the state with

[15] Remember that we know that the phase boundary has to extend into the $\lambda > 0$ region since the effect of small λ is just to slighly renormalize the gauge coupling; this has the effect of shifting the value of the microscopic parameter g at which the transition occurs, but cannot eliminate the existence of the transition. Hence some sort of order parameter should remain well-defined upon adding in nonzero λ.

two charges separated by a distance L. The division by the closed Wilson loop is just to normalize the two-charge state correctly.

If $C(L) = 0$ as $L \to \infty$, then the two-charge state is orthogonal to the vacuum $|0\rangle$. Since any state can be expanded in terms of the spectrum of the Hamiltonian, and since $|0\rangle$ is in the spectrum, $C(L) = 0$ means that the two-charge state $|\bullet\,\bullet\rangle$ is part of the excitation spectrum of the theory. Since we have taken $L \to \infty$, this means that the spectrum of the theory includes free isolated charges—this is what we expect from the deconfined / topological phase.

On the other hand if $C(L) \neq 0$ in the $L \to \infty$ limit, then $|\bullet\bullet\rangle$ cannot be in the excitation spectrum, since it has overlap with the ground state. In this case free charges don't appear in the spectrum, which is what we expect from the confined phase.

This expectations are indeed borne out, at least at small λ where we can calculate the dominant L behavior of $C(L)$. First of all, we know from earlier that for large L, we have asymptotically (redefining g^2 by a factor of 4 so that the exponent is prettier)

$$\langle \prod_{l\in\square} \mathcal{Z}_l \rangle \sim \begin{cases} e^{-4L\ln(\lambda^{-1})} & \text{confined} \\ e^{-4Lg^2} & \text{deconfined} \end{cases} \tag{27}$$

Now for the $\sqcup$ part. The two Z_i, Z_{i+L} variables mean that the first term to survive in a small-λ expansion of the action is a string of $Z^\dagger \mathcal{Z} Z$s connecting the two $\sqcup$ endpoints. In the confined phase, and to lowest order in λ, this string will follow the path of the $\sqcup$; that way it completely screens the electric flux and gives a nonzero contribution to the expectation value at order λ^{2L} ($2L$ is the length of the $\sqcup$). In the deconfined phase, the first term appears at lower order in λ: we can connect the endpoints of the $\sqcup$ with a straight line of $Z^\dagger \mathcal{Z} Z$s at order λ^L, and then add in a contribution of e^{-2Lg^2} from the line tension of the electric flux line. This is larger than screening the whole line with electric charges provided that $\lambda < g$, which we will assume. So then

$$\langle Z_i^\dagger \left(\prod_{l\in\sqcup} \mathcal{Z}_l \right) Z_{i+L} \rangle = \begin{cases} e^{-2L\ln(\lambda^{-1})} & \text{confined} \\ e^{-L\ln(\lambda^{-1}) - 2Lg^2} & \text{deconfined.} \end{cases} \tag{28}$$

We then see that

$$C(L) \sim \begin{cases} \# > 0 & \text{confined} \\ e^{-L\ln(\lambda^{-1})} \to 0 & \text{deconfined,} \end{cases} \tag{29}$$

where # is some number that does not vanish as $L \to \infty$ (literally using our formulae this number would be 1, but we have only been calculating the leading L dependence; in reality $0 < \# < 1$ as $L \to \infty$). This then confirms that $C(L)$ serves as an OP to distinguish the two phases even when $\lambda > 0$ (at least, for perturbatively small λ).

Symmetries

Finally let's talk about the symmetries in the pure gauge theory. With no electric matter we of course have a $\mathbb{Z}_N^{(1)}$ or $U(1)^{(1)}$ symmetry. Somewhat confusing is the fact that the magnetic symmetry is different in the two cases: in the $\mathbb{Z}_N$ case it is a $\mathbb{Z}_N^{(D-2)}$ symmetry (which on the lattice is only ever just an emergent symmetry), while for $U(1)$ it is a $U(1)^{(D-3)}$ symmetry (if it exists).[16] In fact, while both these symmetries have a magnetic character, their origins are pretty different, and it is probably best to not discuss them in the same context.

In discrete gauge theories, the magnetic $\mathbb{Z}_N^{(D-2)}$ symmetry comes emerges in the IR when we work below the gap of the $B_\square$ term. At these low energies, we can work in a constrained subspace where the t' Hooft operators

$$T_\Sigma = \prod_{l \pitchfork \Sigma} \mathcal{X}_l, \qquad \Sigma \in C_{D-2}(\Lambda^\vee; \mathbb{Z}) \tag{30}$$

can't end (that is, they can only be defined for Σ with $\partial\Sigma = 0$), because they create energetically costly magnetic flux along $\partial\Sigma$. Therefore the conserved objects responsible for the $\mathbb{Z}_N^{(D-2)}$ symmetry are the lines / surfaces of the t' Hooft operators, which being $(D-2)$-dimensional give us a $\mathbb{Z}_N^{(D-2)}$ symmetry.

On the other hand, in $U(1)$ gauge theory t' Hooft operators can always be defined on manifolds with $\partial\Sigma \neq 0$, regardless of the value of g: this is because they can be used to create arbitrarily small amounts of magnetic flux along $\partial\Sigma$, which incurs arbitrarily small energetic cost. Therefore t Hooft operators for $U(1)$ are always breakable, unlike for $\mathbb{Z}_N$.

Instead, the magnetic symmetry for $U(1)$ comes from something totally different—the absence of magnetic monopoles. Since the operator that counts monopoles is always two-dimensional, this symmetry is indeed a $(D-3)$-form symmetry. This symmetry and the absence of monopoles is something that's well-defined in the continuum limit.

However, on the lattice I think this symmetry is always non-existent, at least if we use the standard Wilsonian form of the action. Indeed, I don't even think that the concept of a magnetic monopole on the lattice is really meaningful / useful, except at long distances and in the $g \to 0$ limit where we can use the weak-coupling continuum action anyway, and do away with the lattice.[17] The naive way of defining monopoles on the lattice would be to sum the "field strengths" $\mathcal{H}_\square$ around the $\cube$. Since modding by 2π is taken into account in the definition of $\mathcal{H}_\square$, this sum indeed produces something in $2\pi\mathbb{Z}$. However, with this definition monopoles may be created at arbitrarily small energy cost, since we may have e.g. one $\square$ contributing a holonomy of $2\pi - \varepsilon$ and another contributing ε, for an energy that vanishes as ε^2. In this sense monopoles are not really well-defined charges, since the field configurations they set up can be arbitrarily close to the vacuum as far as energetics goes (as long as the magnetic flux is collimated along a certain direction).

Part of this problem is that assigning $\mathcal{H}_\square$ to the field strength at $\square$ is only a reasonable thing to do in the continuum limit where all $\mathcal{H}_\square$s are small. More field strength should mean more energy, which is not true if we assign field strength to $\mathcal{H}_\square$ because of the cosine. Instead, field strength can be defined through the analogue of the continuum equation $d^\dagger F = J$, which is obtained by varying the action with respect to the A_ls. When we do this on the Wilsonian

[16] Mathematically, the shift in the degrees comes from the fact that the group cohomology satisfies $H^k(G; U(1)) \cong H^{k+1}(G; \mathbb{Z})$ as derivable from $1 \to \mathbb{Z} \to \mathbb{R} \to U(1) \to 1$.

[17] $g \to 0$ isn't the only way to take the continuum limit, since there's a critical point at nonzero g_* for the lattice theory. Here the action we should write down is unknown (at least to me), and there's no obvious way of finding the continuum theory from the lattice model. Presumably though the holonomies are not slowly varying on the scale of a lattice, so that expanding the cosine is not allowed.

action we get

$$d^\dagger \sin \mathcal{H}_\Box = J \implies F_\Box = \sin \mathcal{H}_\Box, \tag{31}$$

where $d^\dagger$ is the lattice divergence. $\sin \mathcal{H}_\Box$ is a reasonable field strength since it vanishes for holonomies which cost zero energy and reduces to $\mathcal{H}_\Box$ in the continuum limit of small $\mathcal{H}_\Box$. One could then define the monopole charge by doing $\sum_{\Box \in \Sigma} \sin \mathcal{H}_\Box$ as a tentative monopole charge. While this is not quantized, it reduces to the usual monopole charge in the continuum limit, and in my opinion better encapsulates the energetical / dynamical properties of magnetic monopoles for all g.

Anyway, the point of the above rambling discussion is that in the weak coupling continuum limit fixed point, we can consider a regime where the monopoles are massive and a $U(1)^{D-3}$ symmetry emerges. Away from $g \to 0$ though this symmetry is explicitly broken as the monopoles start to condense upon approaching the confinement phase transition.

2 Transfer matrices and quantum-to-classical heuristics

The quantum-to-classical mapping for lattice models a la Kogut used to seem rather ad-hoc to me—in today's entry we'll make a few comments that made me feel a little more confident about the whole procedure.

Coupled springs

A good first example is a 1d classical model of springs:

$$Z = \int \prod_i dr_i \, e^{-\frac{1}{2K}\sum_i (r_i - r_{i+1})^2}, \tag{32}$$

where we e.g. fix ∂conds to be $r_0 = r_I$ and $r_N = r_F$, or else work with periodic ∂ conds. The goal is to rewrite this as a quantum problem in $0+1$D. This entails writing the partition function as $\langle r_F | e^{-\beta_Q H_Q} | r_I \rangle$ for some parameter β_Q and some 0d Hamiltonian H_Q. Therefore we need to collapse the product over dr_is into a trace of a power of a single matrix.

The key to doing this is to realize that for any function $f(r)$ (that can be Fourier transformed), we have

$$e^{K\nabla^2/2} f(r) = \int dr' \, e^{-(r-r')^2/2K} f(r'). \tag{33}$$

To prove this, it is enough to consider the case where $f(r) = e^{iqr}$, in which case

$$\int dr'\, e^{-(r-r')^2/2K} e^{iqr'} = e^{iqr} \int dr'\, e^{-(r')^2/2K+iqr'} = e^{iqr} e^{-q^2K/2} = e^{K\nabla^2/2}(e^{iqr}). \tag{34}$$

The sign of the exponent can be remembered by recalling that ∇^2 is negative-semidefinite.

Therefore the matrix $e^{K\nabla_r^2/2}$ is a "contraction matrix" that does the needed integral in the partition function over the coordinate r. Hence we can write the partition function as

$$Z = \langle r_F | \left(e^{K\nabla^2/2} \right)^N | r_I \rangle. \tag{35}$$

Evidently this represents time evolution by the operator

$$\beta_Q H_Q(p) = \frac{NK}{2} p^2, \tag{36}$$

where now $p = -i\nabla$. So indeed, we can rewrite the partition function as time evolution for a 0+1D quantum problem.

It is then straightforward to port this concept of transfer matrices up to higher dimensional versions of the spring model. Let us consider the classical model with Hamiltonian

$$\beta H_C = \frac{1}{2K_\perp} \sum_{\langle ij \rangle_\perp} (r_i - r_j)^2 + \frac{1}{2K_\parallel} \sum_{\langle ij \rangle_\parallel} (r_i - r_j)^2. \tag{37}$$

Here $\langle ij \rangle_\perp$ are all the links perpendicular to a certain distinguished axis (which will become time), and $\langle ij \rangle_\parallel$ are all the links parallel to the same axis. The matrix that does the contracting between different "time slices" is the matrix that contracts the $\perp$ hyperplanes together by doing the integrals over the $\parallel$ links. We see that this "contraction matrix" is just $C = \exp\left(\sum_i K_\parallel \nabla_i^2/2\right)$, and that it acts on the matrix $M = \sum_{\langle ij \rangle} (2K_\perp)^{-1}(r_i - r_j)^2$, where the sum over nns only occurs within one hyperplane. Therefore the integral over all coordinates in a given hyperplane is accomplished with the matrix CM, and so

$$Z = \langle \partial_F | (CM)^N | \partial_I \rangle, \tag{38}$$

where again N is the (dimensionless) length of the "time" direction and ∂_I, ∂_F are boundary conditions. Then if we have PBC in the $\parallel$ direction this gives us a quantum problem with the Hamiltonian

$$H_Q \propto -\frac{K_\parallel}{2} \sum_i \nabla_i^2 + \frac{1}{2K_\perp} \sum_{\langle ij \rangle} (r_i - r_j)^2, \tag{39}$$

which is exactly what we expect. The general point is that the contraction over the time direction can be implemented by new non-commuting momentum variables, and when this is done we get a quantum Hamiltonian.

Actually we've cheated a bit here—since ∇^2 and $(r_i - r_j)^2$ don't commute, we can't combine the C and M matrices into a single exponential. We can only do this if the commutator of the logs of C and M is small enough to be dropped in the exponential. This means that we need the limit[18]

$$\frac{K_\parallel}{K_\perp} \to 0. \tag{40}$$

Noting that the Ks have dimensions of length, we interpret this to mean that the quantum partition function is strictly speaking only obtained in the limit where the lattice spacing in the time ($\parallel$) direction is taken to be much smaller than the spacing in the spatial directions. That is, the theory becomes that of a quantum system only when we take the continuum limit in the time direction.[19] If we write the number of lattice spacings in the $\parallel$ direction as $N = 1/Ta$ where T^{-1} and a have dimensions of "time",

$$Z = \mathrm{Tr}[e^{-T^{-1}H_Q}], \qquad -H_Q = \frac{K_\parallel}{2a}\sum_i \nabla_i^2 + \frac{K_\perp}{2a}\sum_{\langle ij\rangle}(r_i - r_j)^2, \tag{42}$$

which is what we expect.

The Ising model

In the case of e.g. the Ising model, we use the same procedure, by introducing operators which don't commute with the spins to write the spin sum as a matrix product. In two dimensions with uniform couplings (for simplicity of notation; generalizing is easy), we start with $H_C \propto \sum_{\langle ij\rangle} Z_i Z_j$. The needed contraction identity is

$$\mathrm{Tr}_j\left[e^{-JZ_iZ_j}f(Z_j)\right]|i\rangle = e^{\alpha X_i}f(Z_i)|i\rangle, \tag{43}$$

where we will find α shortly, and where $|i\rangle$ can be either $|\uparrow\rangle$ or $|\downarrow\rangle$. Writing $f(Z) = f_e + f_o Z$, the sum is, after some algebra,

$$\mathrm{Tr}_j\left[e^{-JZ_iZ_j}f(Z_j)\right]|i\rangle = \left[\begin{pmatrix}\zeta^{-1} & \zeta \\ \zeta & \zeta^{-1}\end{pmatrix}\begin{pmatrix}f_e + f_o \\ f_e - f_o\end{pmatrix}\right]_i = \left[(\mathbf{1}\zeta^{-1} + X\zeta)\begin{pmatrix}f_e + f_o \\ f_e - f_o\end{pmatrix}\right]_i \tag{44}$$

where $\zeta \equiv e^J$. Therefore we need

$$\cosh\alpha = \zeta^{-1}, \qquad \sinh\alpha = \zeta \implies \alpha = \tanh^{-1}[\zeta^2]. \tag{45}$$

Each 1d line in the classical partition function is then integrated out with the matrix CM, where now $C = e^{\tanh^{-1}[\zeta^2]\sum_i X_i}$ and $M = e^{-J\sum_{\langle ij\rangle} Z_iZ_j}$, which lets us write Z as time evolution with the 1d TFIM Hamiltonian

$$H_Q \propto J\sum_{\langle ij\rangle} Z_iZ_j - \tanh^{-1}[\zeta^2]\sum_i X_i. \tag{46}$$

[18] Really, we want this ratio to be small compared to the typical eigenvalues of the commutator of the ∇^2 and $(r_i - r_j)^2$ terms that appear in the trace.

[19] Strictly speaking we can always write the classical partition function as a quantum system with Hamiltonain

$$T^{-1}H_{\mathrm{awful}} = -N\ln(CM), \tag{41}$$

where T^{-1} is the temperature of the quantum system. But this expression is horrendous, and certainly not something we want to map our classical model to, unless the assumptions on the Ks hold.

$U(1)$ gauge theory

Now let's do $U(1)$ gauge theory. We will start with the quantum theory and map to the classical one, for reasons that will become clear after we get the answer. Therefore we start with (probably should have relabeled g as g^2; oh well)

$$H = -\frac{1}{2g}\sum_{\square}\mathcal{F}_{\square}(U_{l\in\partial\square}) - \frac{g}{2}\sum_{l} E_l^2. \tag{47}$$

Here $\mathcal{F}_\square$ is some function of the link variables on $\square$—it could e.g. be $\cos(\mathcal{H}_\square)$ where $\mathcal{H}_\square$ is the holonomy around $\square$, but we will leave it unspecified for now. We will do the $\mathcal{QC}$ mapping first (which given the usual Trotterization procedure is essentially unambiguously defined), and then see what types of plaquette terms are generated on the new temporal plaquettes—this will inform us about which form of $\mathcal{H}_\square$ is "best" to work with.

Again we will be splitting up into "time" steps of size $\delta^{-1} \gg E_{\text{typical}}$. At each step, we will need to insert the gauge invariance projector, which is

$$\Pi = \prod_v \int \mathcal{D}\theta_v^0\, e^{i\theta\sum_{l\in\partial v} E_l}, \tag{48}$$

where θ_v^0 will be thought of as proportional to the logaraithm of a link variable U_l living on a link extending out of v and up into the "time" direction. Inserting Π at each timestep is needed since in the resolutions of $\mathbf{1}$ that are inserted at each step, we will be summing over *all* configurations for the link variables, not just the ones that obey Gauss's law.[20] One step in the Trotterized partition function, giving the matrix element between slices at timesteps j and $j+1$, is then

$$\prod_{v.l}\int \mathcal{D}E_l\,\mathcal{D}\theta_v^0\,\langle U_{j+1}|E\rangle\langle E|U_j\rangle \exp\left(\frac{g\delta}{2}\sum_l E_l^2 + \frac{\delta}{2g}\sum_\square K(U_{l\in\partial\square}) + i\theta_v^0\sum_{l\in\partial v}E_l\right). \tag{49}$$

In order to write down the above, we had to make an assumption: if $\langle E|e^{-\delta H}|U_j\rangle$ is to be simple, we need to split up the E and U parts of H, so that they can act on their respective eigenstates. One assumption that allows us to do this (but this may be slightly more strong than we need) is to assume that $\delta H \ll 1$, by which we mean that $\delta E \ll 1$ for any typical energy E that makes a non-negligible contribution to the path integral. [21] We will be sticking with δ small enough that this assumption is valid, which is equivalent to taking the continuum limit in the time direction.

Now we want to do the sum (despite the notation it's a sum, since the spectrum of the Es is $\mathbb{Z}$) over all values of the E_ls. We do this with Poisson resummation: for any function $f(x)\in L^1(\mathbb{R})$, the sums of f over $\mathbb{Z}$ and the Fourier transform $\widetilde{f}$ over $2\pi\mathbb{Z}$ are equal:

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{p\in\mathbb{Z}}\widetilde{f}(2\pi p). \tag{50}$$

[20]We want to sum over all configurations since matrix elements like $\langle U_j|E\rangle$ are easier to deal with when the sum is over all U_js and Es, not just those that obey Gauss's law.

[21]Since H is unbounded we really need $\delta = 0$, but presumably the super high eigenstates of H aren't all that important for the questions we want to know. We will have in mind taking $\delta < \Lambda^{-1}$ where Λ is some finite but "large" cutoff scale.

A special application of this is to the function $f(x) = e^{-\frac{1}{2}x^2+ix\theta}$, for which we get

$$\sum_n e^{-\alpha n^2/2} e^{in\theta} = \sqrt{\pi/\alpha} \sum_p e^{-\frac{1}{2\alpha}(\theta-2\pi p)^2}. \tag{51}$$

Now since E_l is conjugate to the log of the U_j link variables, we have $\langle E_l | U_l \rangle = e^{iE_l A_l}$, where $U_l = e^{iA_l}$.[22] Each E_l therefore appears in the exponent as

$$g\delta E_l^2 + iE_l(A_{l;j} - A_{l;j+1} + \theta^0_{vL} - \theta^0_{vR}), \tag{52}$$

where $\theta^0_{vL/R}$ are the Lagrange multiplier variables at the end / beginning of the link L. For obvious reasons, let us denote the quantity in parenthesis as $\mathcal{H}_{\square_t}$. Then using the Poisson summation formula, and taking the product over all timesteps, the partition function becomes

$$Z = \sum_{\{p\}} \int \prod_l \mathcal{D}U_l \prod_v \mathcal{D}\theta^0_v \exp\left(-\frac{1}{2g\delta} \sum_{\square_t} (\mathcal{H}_{\square_t} - 2\pi p_{\square_t})^2 - \frac{\delta}{2g} \sum_{\square} \mathcal{F}_\square(U_{l\in\partial\square}) \right), \tag{53}$$

where $\square_t$ are the temporal plaquettes and l the spatial links. Now the usual Wilsonian action would have a $\cos(\mathcal{H}_{\square_t})$ instead of the sum over all ps. This shows us that the Wilsonian lattice action is slightly less "natural", in the sense that it is not what we get when we do the above $\mathcal{QC}$ mapping on a quantum Hamiltonain (regardless of what we choose for $\mathcal{F}_\square$). To make this symmetric in spacetime then (despite the fact that the image of the $\mathcal{QC}$ mapping is very asymmetric in the time direction because of the annoying δ), it is nice to similarly choose the $\mathcal{F}_\square$ spatial part of the action to have the same Villain form. Therefore the classical partition function we get is, now writing the θ^0_vs as link variables on a $D+0$ dimensional spacetime and writing the couplings in plane and along the time direction as $K_s = \delta/2g$ and $K_t = 1/(2\delta g)$ respectively,

$$Z = \sum_{\{p\}} \sum_{\{z_l\}} \exp\left(-K_t \sum_{\square_t} (\mathcal{H}_{\square_t} - 2\pi p_{\square_t})^2 - K_s \sum_{\square} (\mathcal{H}_\square - 2\pi p_\square)^2 \right). \tag{54}$$

The usual Wilsonian cosine form is recovered in the limit where both the K_t and K_s couplings are small, so that we may use the Villain approximation

$$e^{\frac{1}{2g}\cos(\theta)} \underset{\sim}{\propto} \sum_{k\in\mathbb{Z}} e^{-\frac{1}{2g}(\theta-2\pi k)^2}, \tag{55}$$

where we don't care about getting the normalization factors exactly right. This approximation gets better as θ gets forced to lie in $2\pi\mathbb{Z}$, i.e. when we are at weak coupling.[23]

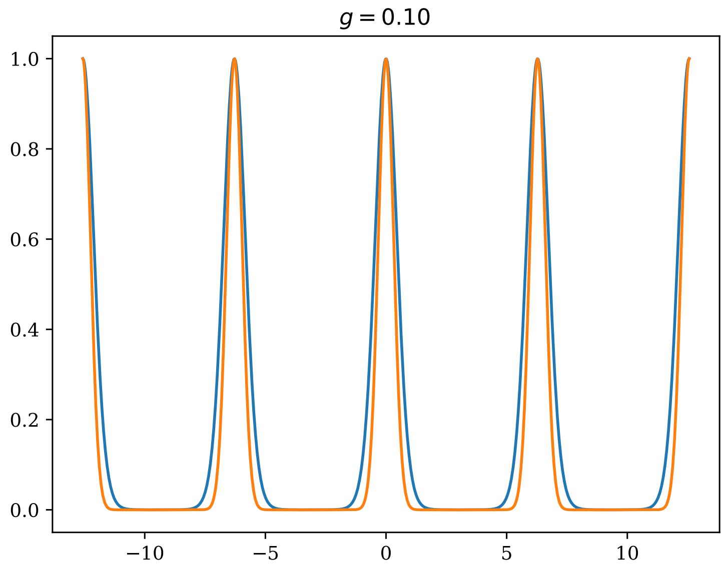
This has been for pure gauge theory, but adding (bosonic, for simplicitly) matter is straightforward. The Wilsonian way to do this would be to add the minmial coupling term

[22] The state $|U_j\rangle$ is the whole state at timestep j, so $|U_j\rangle = \bigotimes_l |U_l\rangle$. This is a bigger $\mathcal{H}$ space than the physical one, but that's okay, since we've added the Πs in between each resolution of the identity in the Trotterization.

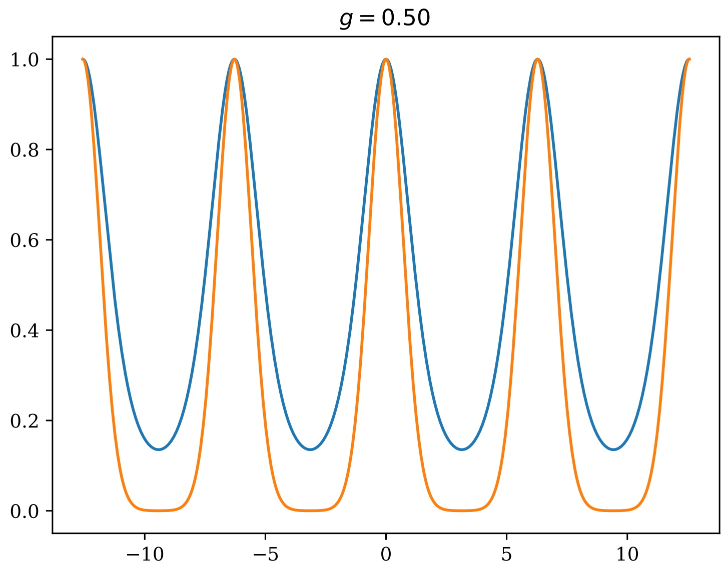
[23] Aside: how well does the Villain approximation work? Let's plot the two functions and find out: for

$\lambda|\psi_i - U_{\langle ij\rangle}\psi_j|^2/2$. The manipulations become simpler if the magnitude of the matter field is frozen out as is the case near the Higgs phase. This simplifies things since then the only dynamical matter is the part of ψ directly responsible for the $U(1)$ symmetry, viz the phase ϕ of ψ. This in turn means that the expression for the generator of gauge transformations G is particularly simple, since the canonical momentum of ϕ is equal to the density appearing

several values of g,



(56)



(57)

in G. We therefore will take the matter part of the Hamiltonian to have the villain form

$$H \supset \frac{\lambda}{2} \sum_l ((d\phi)_l - A_l - 2\pi m_l)^2 + \frac{1}{2\lambda} \sum_v \pi_\phi^2 \tag{59}$$

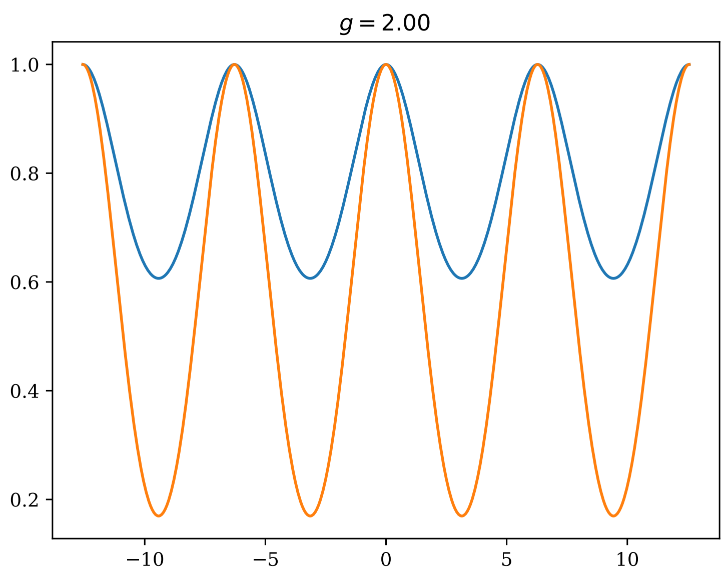
where $m_l \in \mathbb{Z}$ is summed over in the path integral and d is the lattice derivative. The projectors that get inserted at each timestep are now

$$\Pi = \prod_v \int \mathcal{D}\theta_v^0 \exp\left[i\theta_v^0 \left(\sum_{l \in \partial v} E_l - \pi_v \right) \right]. \tag{60}$$

We then Poisson-resum the integers appearing in the kinetic terms for the matter fields and gauge fields. The summations over E_l and π_ϕ are identical in form (since π_ϕ is the momentum for a phase variable, it too is valued in $\mathbb{Z}$), and we end up with the classical partition function

$$Z = \sum_{\{p\},\{q\}} \sum_{\{z_l\},\{\phi_v\}} \exp\left(-S_{\text{gauge}} - \frac{\lambda}{2\delta} \sum_{l_t} ((d\phi)_{l_t} - A_{l_t} - 2\pi q_{l_t})^2 - \frac{\lambda\delta}{2} \sum_l ((d\phi_l) - A_l - 2\pi q_l)^2 \right), \tag{61}$$

with S_{gauge} the action appearing in the pure gauge partiton function above and l_t, l are temporal and spatial links, respectively.



(58)

So even when $g = 2$ it's not unreasonable to assume that the universal behaviors of models with the typical $\cos(B_\square)$ action will still be in the same universality class as the "exact" model coming from the quantum-to-classical mapping.

$\mathbb{Z}_N$ gauge theory

As a more complicated example, consider quantum $\mathbb{Z}_N$ lattice gauge theory. One way to write the action is in the Wilsonian form

$$H = -\frac{1}{2g}\sum_{\square} B_{\square} - \frac{g}{2}\sum_l X_l + h.c. \tag{62}$$

where $ZX = \zeta_n XZ$. Again it's a bit easier to go from quantum to classical here, so that's what we'll do.[24] At a technical level, the $\mathbb{Z}_2$ case is a bit different from the $N > 2$ case. We will first treat the $\mathbb{Z}_2$ case, using the above Wilsonian form of the action, and then come back to the more general case.

When we trotterize we do the usual thing of inserting both position (Z basis) and momentum (X basis) resolutions of the identity. Again, let the time step interval be δ, so that the inverse temperature in the quantum model is $\beta = N\delta$. Then

$$\langle X|\Pi e^{-\delta H}|Z\rangle \approx \frac{1}{2}\prod_v\left(1+\prod_{l\in\partial v} X_l\right) e^{\delta g^{-1}\sum_\square B_\square(Z)+\delta g\sum_l X_l}\langle Z|X\rangle, \tag{63}$$

where now the Xs and Zs are supposed to be thought of as numbers, not matrices. Now $\langle Z|X\rangle$ gives the matrix elements of the discrete Fourier transform, namely

$$\langle Z|X\rangle = \zeta_2^{zx}, \tag{64}$$

where we will use lowercase letters to denote the elements of $\mathbb{Z}_N$ with the group law being addition, i.e. $z, x \in \mathbb{Z}_N \subset \mathbb{N}$ (here $\mathbb{Z}_2$). This means that e.g. $z = N(\ln Z)/(2\pi i)$. We can then write the gauge invariance projector with a Lagrange multiplier as

$$\Pi = \frac{1}{2}\prod_v\sum_{k_v}\zeta_2^{k\sum_{l\in\partial v} x_l}. \tag{65}$$

Finally the X term in the Hamiltonian can be re-written as $\delta g(1-2x)$, so that the timestep is (dropping constants)

$$Z_{j\to j+1} = \prod_{l,v}\sum_{x_l,k_v}\zeta_2^{\sum_l x_l(z_{l;j}-z_{l;j+1}+k_{v0}-k_{v1})}\exp\left(\delta g^{-1}\sum_\square B_\square(z_j)+\delta g\sum_l(1-2x_l)\right). \tag{66}$$

Now we need to do the sum over the x_ls. Because we are only summing over $\mathbb{Z}_2$, this is easy to do: the relevant sum gives

$$e^{\delta g} + e^{-\delta g}\zeta_2^{\mathcal{H}_{\square_t}(z,k)}, \tag{67}$$

where $\mathcal{H}_{\square_t} \in \mathbb{Z}$ is the obvious notation for the discrete holonomy around the temporal plaquette $\square_t$. Since $e^{az} = \cosh a + s\sinh a$ if $s \in \pm 1$, we can define a new coupling g_t^{-1} in terms of which the sum is

$$e^{\delta g} + e^{-\delta g}\zeta_2^{\mathcal{H}_{\square_t}(z,k)} = \exp\left(g_t^{-1}\delta\zeta_2^{\mathcal{H}_{\square_t}(z,k)}\right) \tag{68}$$

[24]In general quantum $\to$ classical is easier than classical $\to$ quantum. This is because when doing the latter there is a canonical procedure for Trotterizing, inserting resolutions of $\mathbf{1}$ in the momentum basis, etc. To do classical $\to$ quantum we have to make smart guesses about how to do the contraction in the transfer matrix, which is harder.

where

$$g_t^{-1}\delta = \tanh^{-1}[e^{-2g\delta}]. \tag{69}$$

Doing the division another way gives instead $g_t^{-1}\delta = -\frac{1}{2}\ln\tanh[g\delta]$, but this is actually self-consistent due to the curious fact that

$$-\frac{1}{2}\ln\tanh\left(-\frac{1}{2}\ln\tanh(s)\right) = s. \tag{70}$$

Now the exponential of $\zeta_2 = -1$ with the $\mathcal{H}$ is just a pretentious way of writing the plaquette operator for the temporal plaquettes, hence we get an anisotropic classical model with partition function

$$Z = \prod_l \sum_{z_l} \exp\left(\frac{\delta}{g}\sum_{\square_s} B_{\square_s}(z) + \frac{\delta}{g_t}\sum_{\square_t} B_{\square_t}(z)\right). \tag{71}$$

As usual, if we hold g fixed then the $\mathcal{QC}$ correspondence is only exact in the limit that the temporal coupling diverges, since $\delta/g_t \to \tanh^{-1}(1) = \infty$.

Now we address the situation for $N > 2$. Things are more complicated now since the $+h.c.$ term in H is nontrivial. Adding in the conjugate gives cosines $\cos(\mathcal{H}_\square), \cos(x_l)$. The latter of these makes doing the sum over x intermediate states in the Trotterization difficult. So like with the $U(1)$ case, we can really only get anywhere if we work in the small-g regime where Villain-ization of the action is appropriate—this also has the technical bonus of allowing us to work directly with the x, z variables rather than the X, Z ones, the former being better suited to doing computations. Anyway we first perform Poisson resummation by doing the sum over m in the term appearing as $(2\pi x_l/N - 2\pi m)^2$. This gives us something linear in x, with x only appearing in the exponents of ζ_Ns, so that x can then be integrated out to give a mod N delta function. This sets $\mathcal{H}_{\square_t} = Np$ for $p \in \mathbb{Z}$, for every temporal plaquette. These manipulations are all basically the same as they were in the $U(1)$ case; when we're done we get

$$Z = \sum_{\{p\}} \int \mathcal{D}U_l \, \exp\left(\frac{\delta}{2g_t}\sum_{\square_t}\left(\frac{2\pi}{N}\mathcal{H}_{\square_t} - 2\pi p_{\square_t}\right)^2 + \frac{\delta}{2g}\sum_{\square}\left(\frac{2\pi}{N}\mathcal{H}_{\square} - 2\pi p_{\square}\right)^2\right), \tag{72}$$

where $g_t^{-1}\delta = \frac{N}{4\pi^2 g\delta}$.

Lattice duality for generalized classical $\mathbb{Z}_N$ spin models with gauge fields

Today we'll derive some dualities involving classical stat mech models of discrete matter + gauge fields on a lattice. Several specific cases of the general formula below appear in various places in the cmt literature, where their derivations take up much more space than is needed. Differential form notation is a huge time-saver!

We will be considering an isotropic D-dimensional stat mech model with partition function

$$Z = \sum_{\substack{m\in C^{q+1}(\Lambda;\mathbb{Z}) \\ n\in C^{q}(\Lambda;\mathbb{Z})}} \sum_{\substack{a\in C^{q}(\Lambda;\mathbb{Z}_N) \\ \phi\in C^{q-1}(\Lambda;\mathbb{Z}_N)}} \exp\left[-\frac{\beta}{2}\left(\frac{2\pi}{N}da - 2\pi m\right)^2 - \frac{\lambda}{2}\left(\frac{2\pi}{N}(d\phi - a) - 2\pi n\right)^2\right], \quad (73)$$

where Λ is some D-dimensional lattice. The sums over m and n can be thought of as coming from making the Villain approximation (large β, λ) to a discrete Wilsonian action written in terms of cosines like $\cos(2\pi da/N)$. Since the cosines are infinite order in the fields but the above representation is quadratic in everything, the Villain form will need to be used to do dualities. The differentials are of course lattice differentials.

The duality works by doing Poisson resummation on the m and n variables. We use

$$\sum_{m\in\mathbb{Z}} \widetilde{f}(2\pi m) = 2\pi \sum_{k\in\mathbb{Z}} f(k) \quad (74)$$

applied to each $q+1$, q cell of Λ. This means that e.g. (dropping proportionality constants)

$$\sum_{m\in C^{q+1}(\Lambda;\mathbb{Z})} \exp\left[-\frac{\beta}{2}\left(\frac{2\pi}{N}da - 2\pi m\right)^2\right] = \sum_{l\in C^{D-q-1}(\Lambda^\vee;\mathbb{Z})} e^{-l^2/2\beta}\zeta_N^{l\wedge da} \quad (75)$$

which we get after doing the fourier transform. Here $\Lambda^\vee$ is the dual lattice; writing l as a $D-q-1$ chain in this way is purely for notation's sake. If we preferred we could take $l \in C^{q+1}(\Lambda;\mathbb{Z})$ and replace the $l\wedge da$ with $\star l \wedge da$.

Doing the same Poisson resummation on the term with the n field, we see that the partition function can equivalently be written as

$$Z = \sum_{\substack{l\in C^{D-q-1}(\Lambda^\vee;\mathbb{Z}) \\ k\in C^{D-q}(\Lambda^\vee;\mathbb{Z})}} \sum_{\substack{a\in C^{q}(\Lambda;\mathbb{Z}_N) \\ \phi\in C^{q-1}(\Lambda;\mathbb{Z}_N)}} \exp\left[-\frac{l^2}{2\beta} - \frac{k^2}{2\lambda} + \frac{2\pi i}{N}\left(l\wedge da + k\wedge(d\phi - a)\right)\right]. \quad (76)$$

Now we can integrate out ϕ and a. Summing over ϕ tells us that

$$k = d\widetilde{a} + N\widetilde{m}, \qquad \widetilde{m}\in C^{D-q}(\Lambda^\vee;\mathbb{Z}), \quad \widetilde{a}\in C^{D-q-1}(\Lambda^\vee;\mathbb{Z}_N). \quad (77)$$

Summing over a then tells us that we can parametrize l as

$$l = d\widetilde{\phi} + \widetilde{a} + N\widetilde{n}, \qquad \widetilde{n}\in C^{D-q-1}(\Lambda^\vee;\mathbb{Z}), \quad \widetilde{\phi}\in C^{D-q-2}(\Lambda^\vee;\mathbb{Z}_N). \quad (78)$$

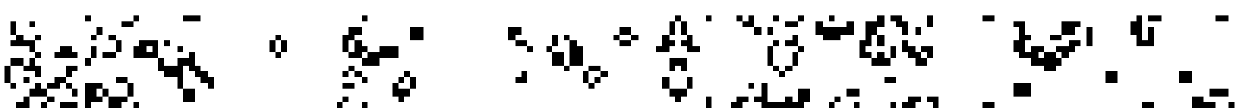
Therefore we can equivalently write the partition function in the dual form

$$Z = \sum_{\substack{\widetilde{m}\in C^{D-q}(\Lambda^\vee;\mathbb{Z}) \\ \widetilde{n}\in C^{D-q-1}(\Lambda^\vee;\mathbb{Z})}} \sum_{\substack{\widetilde{a}\in C^{D-q-1}(\Lambda;\mathbb{Z}_N) \\ \widetilde{\phi}\in C^{D-q-2}(\Lambda;\mathbb{Z}_N)}} \exp\left[-\frac{\widetilde{\beta}}{2}\left(\frac{2\pi}{N}d\widetilde{a} - 2\pi\widetilde{m}\right)^2 - \frac{\widetilde{\lambda}}{2}\left(\frac{2\pi}{N}(d\widetilde{\phi} - \widetilde{a}) - 2\pi\widetilde{n}\right)^2\right] \tag{79}$$

where the dual couplings are defined as

$$\widetilde{\beta} \equiv \frac{N^2}{\lambda(2\pi)^2}, \qquad \widetilde{\lambda} \equiv \frac{N^2}{\beta(2\pi)^2}. \tag{80}$$

Thus a $\mathbb{Z}_N$ theory of $(q-1)$-form variables coupled to a q-form gauge field with gauge and Higgs couplings (β, λ)[25] is the same as that of a theory of $(D-q-2)$-form matter on the dual lattice, coupled to a $(D-q-1)$-form gauge field, with couplings $(N^2/4\pi^2\lambda, N^2/4\pi^2\beta)$. The most useful examples of this are the strong-weak self-duality of gauge fields and matter in three dimensions, and the self-duality of the pure gauge theory in four dimensions, both of which of course have counterparts in the 2+1D and 3+1D quantum models.



References

[1] D. A. Huse and S. Leibler. Are sponge phases of membranes experimental gauge-higgs systems? *Physical review letters*, 66(4):437, 1991.

[25] Okay, β is more like the inverse of the square of the gauge coupling.