Problem Sheet 10

I. PROBLEMS RELATED TO THE LAST LECTURE WEEK: SPIN MODELS WITH CONTINUOUS SYMMETRY IN 2D: KOSTERLITZ THOULESS TRANSITION

A. XY model at low T

Consider the XY-model on a two-dimensional lattice. On each latice site there is a (classical) spin of length $|\vec{S}| = 1$, $\vec{S}_i = (S_i^x, S_i^y) = (\cos(\theta), \sin(\theta))$ which lies in the XY-plane.

The classical Hamiltonian is

$$H = -J \sum_{<i,j>} \vec{S}_i \cdot \vec{S}_j = -J \sum_{<i,j>} \cos(\theta_i - \theta_j). \tag{1.1}$$

Compute (like in the lecture) the spin-spin correlation function

$$< \vec{S}_i \cdot \vec{S}_j > \tag{1.2}$$

at low temperature as a function of the distance $r \equiv r_i - r_j$. Use that at low T you can approximate the Hamiltonian as

$$H = const. + \frac{J}{2} \sum_{<i,j>} (\theta_i - \theta_j)^2. \tag{1.3}$$

[Motivate this approximation!]

Recipe for the calculation: Go to the Fourier transform of θ_i. Show that because θ is real one has $\theta(k) = \overline{\theta(-k)}$. Write the Hamiltonian in terms of $\theta(k)$ and conclude that the real and imaginary part of $\theta(k)$ are independent Gaussian variables, and determine their mean and variance (as in the lecture - however, notice that the Hamiltonian in the form (1.3) is still a lattice version, which gives a slightly more involved form of the variance of $\mathrm{Var}(Re\theta_k) = \mathrm{Var}(Im\theta_k)$ than what we did in the lecture). Express $\vec{S}_i \cdot \vec{S}_j = \mathrm{Re}(e^{i(\theta_i - \theta_j)})$ in terms of Fourier components and compute the Gaussian average! You end up with a sum over the independent $Re(\theta_k)$ and $Im(\theta_k)$. Since the interesting physics comes from longwavelengths, i.e., small k, you can take the continuum limit and assume $ak \ll 1$, i.e. you may approximate $\cos(ak) \approx 1 - (ak)^2/2$. You get an exponent of an integral over k. The remaining integral over k is difficult to do exactly, but you can extract the leading logarithmic behavior for large r/a (up to constant which is a bit more difficult to determine).

With the above ingredients, find the power law decay of the spin-spin correlator! The final result is

$$< \vec{S}_i \cdot \vec{S}_j >= C \left(\frac{a}{|r_i - r_j|} \right)^{T/4\pi J} \tag{1.4}$$

where C is a constant, and the result is valid for $|r_i - r_j| \gg a$.

This is the characteristic and unusual "critical" decay of correlations in the low T phase of systems with continuous symmetry in 2d. Of what nature is the decay in 1d instead?

Remark: 2d is special because it is the limiting (or, marginal) case where the Mermin-Wagner theorem (forbidding true long range order) just still applies. For this reason 2d is called the lower critical dimension of the models considered. (Instead for the Ising model, the lower critical dimension is 1d.) In such limiting situations one usually finds 'marginal' behavior (logarithms instead of power laws, or power laws instead of exponentials).

Explain what one means by: 'The spin-spin correlations in 2d are critical and scale-free"?

Argue why you would indeed expect that the exponent of the power law (1.4) decreases with decreasing T.

B. XY model at high T

Without making the approximation (1.3), consider the original XY model at high temperatures. Let us try to find the spin-spin correlator with a high temperature expansion, which is justified when $T \gg J$. To this end write

$$< \vec{S}_i \cdot \vec{S}_j >= \frac{Tr_{\{S_k\}} \vec{S}_i \cdot \vec{S}_j \exp[-\beta \sum_{<k,l>} \vec{S}_k \cdot \vec{S}_l]}{Tr_{\{S_k\}} \exp[-\beta \sum_{<k,l>} \vec{S}_k \cdot \vec{S}_l]} \tag{1.5}$$

and expand in β. Note that a term which does not contain every spin variable S_i an even number of times will yield zero! Convince yourself that the leading order term in β (having the smallest power of β) will be a term where a chain of nearest neighbors $< k, l >$ connects i and j in the straightest possible manner. Using this insight, evaluate this leading order term (i.e.: compute its spin trace!) and extract from this the leading behavior of the correlation function for two sites with the same y-coordinate (both lying on the x-axis, e.g.).

The final result should be:

$$< \vec{S}_i \cdot \vec{S}_j > \sim \exp(-|r_i - r_j|/\xi) \tag{1.6}$$

with

$$\xi = 1/\log(2/\beta J). \tag{1.7}$$

Even without explicit calculation it should become clear that at high T the correlation must die off exponentially: Each spin can transmit only part of the information about its orientation to its neighbors (because of thermal fluctuations). Thus, at each step leading further away from the site i a roughly equal amount of information gets lost, or in other words the correlations diminish with a constant factor per step on the lattice, and the correlations thus decrease exponentially. This consideration is much more general than the considered spin system. It is the reason why in any system with short range interactions

you see exponentially decaying correlations at high T, as well as, usually, at low T once you subtract the expectation values of the spins.

Now interpret this result: How does the correlation length behave as a function of T? Does it behave as you would expect?

Important remark: We have established two qualitatively different behaviors of $< \vec{S}_i \cdot \vec{S}_j >$ at $T \gg J$ and $T \ll J$: One decays exponentially with large distance, the other one only as a power law. Make it clear to yourself that there must thus be a well-defined temperature T_c where the exponential tail disappears and gives way to a power law behavior. This can only happen at a phase transition! We have thus found indirect evidence that there MUST be a phase transition in the 2d XY model!

Find a crude estimate for T_c (just an order of magnitude).

By the way: a typical exchange energy is of order of $0.01 - 0.1eV$. At what temperatures do you thus expect to typically see magnetic order?

C. Destruction of the quasi long range order QLRO by vortices ("topological defects")

Better insight into the above phase transition comes from looking at certain spin configurations which are not captured by the approximation (1.3). Namely, consider spin configurations

$$\vec{S}_i = (x_i, y_i)/\sqrt{x_i^2 + y_i^2} \tag{1.8}$$

or

$$\vec{S}_i = (-y_i, x_i)/\sqrt{x_i^2 + y_i^2}. \tag{1.9}$$

Draw them! These are called vortex configurations. Note that the angles of the spins vary smoothly at large r, but not at small r. Because of the latter, these configurations are not captured by (1.3). Indeed they are at the root of the disappearance of QLRO at high T (while the approximation (1.3) wrongly predicts QLRO at all T's!).

Look at a closed line circling the origin at a large enough distance from the origin, and consider the variation of the angle of S_i along the line. Convince yourself that the total variation must always be a multiple of 2π, and that this multiple does not depend on the line chosen as long as the angles vary smoothly at large r. This latter fact is the reason why one calls such vortex structures "topological". Determine the multiple of 2π in the above cases. This number is actually sometimes called "topological charge". Structures with negative multiples may be called antivortices.

Determine the extra energy of such a configuration as compared to the ground state (all spins ferromagnetically aligned). Show that it diverges with system size as $\log(L)$ in two dimensions!

On the other hand: What is the entropy for the creation of a single vortex anywhere in the sample?

Estimate the energy cost of a vortex and antivortex pair which are at a distance r from each other. (This means the spin configuration where at each site you add the rotations of the spin with respect to some chosen ferromagnetic direction). Don't make an explicit calculation! Just notice that at large distance the angular pattern becomes ferromagnetic and thus does not contribute extra energy. What is the distance scale where this crossover to a ferromagnetic pattern happens roughly? Argue that such a vortex-antivortex pair costs an energy which grows logarithmically with their mutual distance r.

Now, at low T it costs obviously far too much to produce an unpaired vortex since its energy grows like $\log(L)$, and the entropy is too low. There are thus only vortex-antivortex pairs that can be thermally excited. However, at higher T, the logarithmic energy cost of an unpaired vortex can be compensated by the logarithmic entropy. What critical temperature do you predict from the balance of energy and entropy of a single vortex?

At this T_c the vortex-antivortex pairs start unbinding in the sense that vortices can get arbitrarily far away from antivortices. Since vortices come along with a non-ferromagnetic spin pattern at large distance, their occurrence in unbound configurations destroys the magnetic quasi-long range order. This is the famous Kosterlitz-Thouless transition for which we found indirect evidence above.

Compare what happens in the 2d XY model with the 1d Ising chain. What is the main difference between vortices and domain walls? What is the physical consequence of this difference for the presence or absence of a transition?

In the lecture you learnt - without explanation - that the Heisenberg model does *not* have QLRO in 2d. This traces back to an important difference in the cost and nature of topological excitations in the XY and the Heisenberg model. To understand this, first locate the precise reason why there is a logarithmic energy cost for vortices in the XY model. Does it come from short distances or long distances? Once you understand this, try to find out how you can construct spin configurations in the Heisenberg model with non-ordered pattern at large distances, such that they do not cost a logarithmic energy, however. Given the existence of such configurations argue that order is destroyed at any finite T.

II. SELFCONSISTENT HARMONIC APPROXIMATION FOR THE MELTING TRANSITION

The melting of a solid into a liquid is a complicated problem which is not fully understood. Here we discuss a toy problem based on Bogoliubov's variational principle. A crystal can be described as atoms which are trapped in their mutually created periodic potential. When the thermal or quantum fluctuations of the positions become too big, the solid melts. A heuristic criterion for the occurrence of melting is the Lindemann criterion which stipulates that the melting transition takes place when the mean square displacement $< u^2 >$ is of the order of the square of the lattice spacing a,

$$< u^2 >= c_L^2 a^2. \tag{2.1}$$

where c_L empirically is of order 0.1-0.2.

We want to make this plausible by a simplified computation. For simplicity we consider a 1d chain of atoms of mass m in a regular array with spacing a. Now we concentrate on one atom and imagine that the other atoms create a potential $V(u) = -V_0 \cos(q_0 u)$ for it. Hereby $q_0 = \pi/a$. (More realistically one should add an extra negative well around $u = 0$ to the periodic potential, so a to describe that the lowest potential energy for an atom is to be at its original place $u = 0$.) Nevertheless our highly simplified model can illustrate what happens in the melting.

The Hamiltonian for the single atom is

$$H = \frac{p^2}{2m} + V(u) \tag{2.2}$$

where $p = -i\hbar d/du$ in quantum mechanics, or $p = m\dot{u}$ in classical mechanics. We now analyze his system with Bogoliubov's variational principle with a trial Hamiltonian

$$H_0 = \frac{p^2}{2m} + \frac{K_0}{2} u^2 \tag{2.3}$$

which we can solve exactly. The variational principle tells us to minimize the trial free energy

$$F_{\text{trial}} = F_0 + < H - H_0 >_{H_0} \tag{2.4}$$

with respect to the variational parameter K_0. (Since the variational Hamiltonian is harmonic this is called the self-consistent harmonic approximation SCHA) Compute F_{trial} analytically within the quantum mechanical model.

If we minimize blindly, we will find that $K_0 = 0$ trivially minimizes the free energy which is negative infinity. What is the reason behind this? (Entropy versus energy!)

Obviously this minimum is nonsensical. Indeed, we want to describe a transition from a bound "crystal" state (atom in the well at the origin) at low T, to a delocalized state

(molten state) at higher T. However, our single atom picture does not make much sense if we allow the atom to delocalize over many wells. (Other atoms would delocalize, too, and we cannot distinguish them. Thus we overcount the entropy with our single atom picture if we don't restrict the spatial phase space.) We should thus better let K_0 be bounded from below such that the positional fluctuations $< u^2 >$ remain smaller than, say a^2. Find an explicit expression for this criterion.

Now, plot $F_{\text{trial}}(K_0)$ as a function of K_0 in the allowed range $K_{min} < K_0 < \infty$. What form do you expect at high T and low T? Confirm that at high T there is only one minimum at K_{min} which corresponds to the molten, delocalized state.

At lower T there is a secondary minimum appearing, which eventually becomes the dominant minimum. Interpret this fact in terms of a first order melting transition! How do you identify the equilibrium transition and the spinodal point of the crystal?

The problem can be solved in full glory in principle, but it is easier to make a classical approximation: The Hamiltonian H_0 defines an oscillator frequency ω_0. Assume now in your calculation that $\beta\hbar\omega_0 \ll 1$ always. (Argue that this is justified when the mass m is large enough!) Now find the equations determining the value of K_0 which minimizes F. Determine now the spinodal of the solid, where the minimum disappears, and solve for the mean square displacement at that spinodal point. The answer is very simple:

$$< u^2 >_{spinodal} q_0^2 = 2 \tag{2.5}$$

Convince yourself that this is indeed of Lindemanns form and determine c_L.

The equilibrium melting occurs at lower T. In contrast to the spinodal it is sensitive to our choice for K_{min}. Nevertheless it is interesting to find the equilibrium transition and the mean square displacement just before the transition (with Mathematica, e.g.). The corresponding c_L is quite a bit smaller, and comes closer to the empirical value.

Knowing the T of the transition, what condition do you need to check a posteriori in order to justify the classical approximation?

At low enough T quantum mechanics surely should dominate. What is the qualitative criterion which tells you whether the system is a liquid or a solid at $T = 0$? (Show that the ratio between $\hbar^2 q_0^2/m$ and V_0 decides whether quantum fluctuations melt the crystal or not! Give a direct physical intuition of why this ratio is important.) What can you thus do in order to make a substance melt at $T = 0$? Why is helium usually liquid at $T = 0$ while other substances are not (under usual conditions at least)?

III. LANDAU THEORY

Landau had the simple but very powerful idea that a continuous phase transition (usually) comes along with a local order parameter $\phi(r)$ which spontaneously breaks the symmetry of the system. We were mostly concerned with the magnetization density $\phi(r) = m(r)$ as the local order parameter. Below the phase transition there is long range order in $\phi(r)$, meaning that $< \phi(r)\phi(r') > \overset{|r-r'|}{\to} < \phi >^2 \neq 0$ at large distance.

Landau's natural idea (which connects in a sense to Bragg-Williams theory) is to ask about the free energy of a system which is constrained to have an order parameter density $\phi(r)$ (after averaging, that is, coarsegraining over small cubes). The resulting $F[\phi]$ can be considered as an inhibited system with inhibition parameter ϕ. Close to a continuous phase transitions, where ϕ can be expected to be small, F can thus certainly be expanded in powers of $\phi(r)$ and, to account for spatial correlations, its gradients. The free energy functional $F[\phi]$ must thereby respect the symmetries of the microscopic Hamiltonian. E.g., if H was invariant under $S_i \to -S_i$, then the free energy must also be invariant under the transformation of the magnetization order parameter $\phi \to -\phi$. This excludes odd powers in the expansion of $F[\phi]$.

A. Mean field theory

The most general free energy functional of a scalar, that is, not vectorial order parameter (such as $\phi = m^z$ in the Ising model) is

$$f(\phi) = \int d^d r \left[\frac{m}{2}\phi^2(x) + w\phi^3(x) + u\phi^4(x) + ... + c/2(\nabla\phi))^2 + ... \right] \tag{3.1}$$

Explain that in the case of the Ising model symmetry requires $w = 0$.

In the simplest mean field theory approach one neglects the gradient term and determines the minimum, $< \phi >$ of $f(\phi)$. For a meaningful minimum to exist one must have $u > 0$ (otherwise one has to include higher order terms which assure that f is bounded from below at large ϕ).

The coefficients m, w, u etc are in general T-dependent. However, it is usually sufficient to consider them constant, and only retain the T-dependence of $m(T) = m_0(T/T_c - 1)$, which is approximated as linear around the (mean field) T_c. If a magnetic field is applied, a further term $-h\phi$ adds to the free energy.

In systems where $w = 0$ is not imposed by symmetry, the cubic term will be there in general. Sketch the form of the free energy in this case for high and low T around T_c. Do you recognize these curves? They are typical for first order transitions. Indeed if you replace ϕ by $\rho - \rho_c$ for the van der Waals gas close to the critical point, you will find a Landau free energy of this form (see also Huang's book!).

For the symmetric systems with $w = 0$, calculate how the magnetization evolves with $T_c - T$, and determine the critical exponent β, defined as $< \phi > \propto (T_c - T)^\beta$.

At $T = T_c$ calculate how the magnetization grows with magnetic field. This determines a further critical exponent, $< \phi > (T = T_c) \propto h^\delta$.

Plugging back into $f(\phi)$ the value $< \phi >$ find the mean field free energy as a function of T, and compute the specific heat. Convince yourself that it does not diverge at the transition. One thus says that $c_V \propto |T - T_c|^\alpha$ with the mean field exponent $\alpha = 0$.

A last exponent of interest describes the divergence of the correlation length upon approach to T_c. A diverging correlation length means that all parts of the system communicate with each other and are correlated or interdependent. Usually this only happens at a critical point, while correlations become short ranged again on the ordered side of a transition. Now determine the mean field behavior of the correlation length by computing for instance the (connected) correlation function $< \phi(r)\phi(r') > - < \phi >^2$ including the gradient term in $f(\phi)$. [You want to subtract the product of expectation values, because one is mainly interested in the additional effect of r on r']. Proceed very similarly as for the XY model going to Fourier space and back to real space again! You will find

$$< \phi(r)\phi(r') > - < \phi >^2 \propto \exp[-|r - r'|/\xi]/|r - r'| \tag{3.2}$$

with $\xi \propto |T - T_c|^{-\nu}$ where the correlation length exponent is $\nu = 1/2$. You can obtain this exponent also very simply from comparing the gradient term with the quadratic 'mass' term $m\phi^2$, requiring that they should compete with each other at long enough distances. A typical value of a derivative is estimated by replacing $\nabla \to 1/\xi$, and thus you easily find $1/\xi^2 \sim T_c - T$ (and similarly for the exponents δ and β).

The important thing to retain from this exercise is the following: The form of free energy is generic, up to symmetry and dimensionality constraints. Thus the conclusions (and in particular the critical exponents) will be the same for systems sharing the same d and symmetries (even though mean field theory is not exact). This leads to the concept of universality: The critical behavior (scaling of ξ, c_V, $\phi(h)$ with $T - T_c$ etc. are the same in such systems! The value of T_c, numerical prefactors in the power laws etc. are of course not universal but depend on the microscopic details of the models.

This idea is put on firm footing by the renormalization group method which starts from Landau's point of view, and goes systematically beyond mean field treatments including fluctuations to higher and higher orders.