Concepts in Condensed Matter Physics: Tutorial III

The Mermin-Wagner theorem

1 Introduction

This tutorial focuses on the famous Mermin-Wagner theorem, which states that in one and two dimensions, continuous symmetries cannot be spontaneously broken at finite temperature in systems with sufficiently short-range interactions. This means that an order parameter associated with such a symmetry can obtain an expectation value only for $T = 0$ exactly, i.e., in the exact ground state (and sometimes even not then). The theorem is very universal and applies, for example, to magnets, solids and superfluids. It illustrates the fact that as the number of dimensions is lowered the fluctuations become more important, to the point where they destroy any potential ordering below the critical dimension.

2 Classical XY model ($O(2)$ model / 2-vector model / classical rotor model)

The classical XY model has a 2-component unit vector, $\vec{s}_i = \begin{pmatrix} \cos(\theta_i) & \sin(\theta_i) \end{pmatrix}$, on each site of a d-dimensional lattice. The vectors configuration energy can generically be governed by the Hamiltonian

$$H = -\frac{1}{2}\sum_{i\neq j} J_{i,j}\vec{s}_i \cdot \vec{s}_j - \sum_i \vec{h}_i \cdot \vec{s}_i = -\frac{1}{2}\sum_{i\neq j} J_{i,j}\cos(\theta_i - \theta_j) - \sum_i h_i \cos(\theta_i) \ ,$$

but we will focus on the case of $J_{i,j} = J$ for nearest neighbours i,j and zero otherwise, as well as $h_i = 0$ (since we examine *spontaneous* symmetry breaking), leading to

$$H = -J\sum_{\langle i,j\rangle} \cos(\theta_i - \theta_j) \ . \tag{1}$$

With this Hamiltonian in mind the system is invariant under the transformation $\theta_i \to \theta_i + c$, i.e., under rotations in the XY plane in which the vectors $\vec{s}_i$ point. In other words the system has a $U(1)$ symmetry for rotation around the z-axis. Notice, however, that minimal energy is obtained when all $\cos(\theta_i - \theta_j) = 1$, namely, when all the classical spins (the vectors) are aligned. Therefore, the ground state spontaneously breaks this $U(1)$ symmetry.

One would naively expect this ferromagnetic phase, where the $U(1)$ rotational symmetry is spontaneously broken, to extend to finite temperatures (at least for low enough temperatures). This expectation is motivated by the naive intuition that the physics at zero temperature should not be different from the physics at an infinitesimal temperature (At sufficiently high temperatures, of course, there must be a transition to a disordered phase). In d=3 this is indeed the case – there is a finite critical temperature, $T_C \sim J$, below which all spins, on average, point in the same direction (they may still fluctuate locally).

To quantitatively characterize order in this system we define the correlation function

$$C\left(\vec{r}_i - \vec{r}_j\right) \equiv \langle \vec{s}_i \cdot \vec{s}_j \rangle = \langle e^{i(\theta_i - \theta_j)} \rangle \ , \tag{2}$$

where the average is over a thermal ensemble

$$\langle f\left(\theta_i, \theta_j\right) \rangle = \frac{\int \prod_l d\theta_l f\left(\theta_i, \theta_j\right) e^{-\beta H}}{\int \prod_l d\theta_l e^{-\beta H}} \ . \tag{3}$$

At zero temperature, namely for $\beta \to \infty$, this average reduces to the ground state expectation value. In the ground state all $\theta_i - \theta_j = 0$ (modulus 2π) and so the correlation function is 1. For finite temperature, if the system remains ordered, on average all spins are still oriented in the same direction. We therefore expect the correlation function to, perhaps reduce a little, but remain finite even at long distances. Indeed, $\lim_{|\vec{r}_i - \vec{r}_j| \to \infty} C\left(\vec{r}_i - \vec{r}_j\right) > 0$ implies the system is ordered. On the other hand, if the system is not ordered distant spins become uncorrelated and we expect the correlation function to decay with some length scale – the correlation length

$$\lim_{|\vec{r}_i - \vec{r}_j| \to \infty} C\left(\vec{r}_i - \vec{r}_j\right) \sim \exp\left[-\frac{|\vec{r}_i - \vec{r}_j|}{\xi}\right] \ . \tag{4}$$

To see if the system is ordered, we first assume it is and approximate the Hamiltonian based on this assumption. Then, using the approximate Hamiltonian we calculate the correlation function (2). If the system is indeed ordered, self-consistency requires that it stay non-zero in the large distance limit.

Assuming the system is ordered, the fluctuations between adjacent spins are expected to be small, so we can expand the cosine in (1) to obtain a quadratic Hamiltonian

$$H \approx E_0 + \frac{J}{2} \sum_{\langle i,j \rangle} \left(\theta_i - \theta_j\right)^2 \ . \tag{5}$$

with which calculating the correlation function (2) is possible. However, to make our life even easier and since we are interested in the long distance behavior, we will forget about the short distance scale, namely the lattice constant, and use the continuum limit version of the Hamiltonian [1]

$$H \approx \frac{\mathcal{J}}{2} \int d^d r \left(\vec{\nabla} \theta\left(\mathbf{r}\right)\right)^2 \ , \tag{6}$$

where we have also dropped the constant E_0 and defined $\mathcal{J} = J a^{2-d}$. Another way of phrasing this approximation is that at low enough temperatures the correlation length ξ is expected to be much larger than the lattice spacing. Consequently, the physics that happens on the scale of ξ is not sensitive to the lattice. As we said before, this step is not actually necessary, but it will simplify our calculations without qualitatively changing the results.

To calculate the correlation function we first diagonalize the Hamiltonian by transforming to Fourier space with the following definition

$$\theta\left(\mathbf{k}\right) \equiv \frac{1}{\left(2\pi\right)^{d/2}} \int d^d r e^{-i\mathbf{k}\cdot\mathbf{r}} \theta\left(\mathbf{r}\right) \ , \tag{7a}$$

$$\theta\left(\mathbf{r}\right) = \frac{1}{\left(2\pi\right)^{d/2}} \int d^d k e^{i\mathbf{k}\cdot\mathbf{r}} \theta\left(\mathbf{k}\right) \ , \tag{7b}$$

$$\int d^d r e^{i\left(\mathbf{k}+\mathbf{q}\right)\cdot\mathbf{r}} = \left(2\pi\right)^d \delta^{(d)}\left(\mathbf{k}+\mathbf{q}\right) \ . \tag{7c}$$

Plugging (7b) into the Hamiltonian and using (7c) we find

$$H = -\frac{\mathcal{J}}{2} \frac{1}{\left(2\pi\right)^d} \int d^d r d^d k d^d q e^{i\left(\mathbf{k}+\mathbf{q}\right)\cdot\mathbf{r}} \left(\mathbf{k}\cdot\mathbf{q}\right) \theta\left(\mathbf{k}\right) \theta\left(\mathbf{q}\right) = \frac{\mathcal{J}}{2} \int d^d k \epsilon\left(\mathbf{k}\right) \left|\theta\left(\mathbf{k}\right)\right|^2 \ , \tag{8}$$

with $\epsilon\left(\mathbf{k}\right) = \left|\mathbf{k}\right|^2$, and we have used the fact that $\theta\left(\mathbf{r}\right)$ is real, meaning that $\theta\left(-\mathbf{k}\right) = \theta^*\left(\mathbf{k}\right)$.

[1] Formally, the way to do this is to first transform to k-space where one can easily take the long distance limit by simply expanding around $\vec{k} = 0$. The last step is to then transform back to real space by essentially replacing $\delta\vec{k} \to -i\vec{\nabla}$

We can now immediately write that

$$\langle\theta(\mathbf{k})\,\theta(\mathbf{q})\rangle = \frac{\int\mathcal{D}\theta\theta(\mathbf{k})\,\theta(\mathbf{q})\,e^{-\beta H}}{\int\mathcal{D}\theta e^{-\beta H}} = \frac{\delta^{(d)}(\mathbf{k}+\mathbf{q})}{\beta\mathcal{J}\epsilon(\mathbf{k})} \tag{9}$$

We are finally ready to compute the correlation function (2). Since our theory is Gaussian we can use

$$C(\mathbf{r}) \equiv \langle e^{i[\theta(\mathbf{r})-\theta(0)]}\rangle = e^{-\frac{1}{2}\langle[\theta(\mathbf{r})-\theta(0)]^2\rangle} = e^{\langle\theta(\mathbf{r})\theta(0)\rangle-\langle\theta(0)\theta(0)\rangle}\ , \tag{10}$$

where we have used the spatial invariance of the correlation function.

Let us now calculate $\langle\theta(\mathbf{r})\,\theta(0)\rangle$,

$$\begin{aligned}\langle\theta(\mathbf{r})\,\theta(0)\rangle &= \int\frac{d^dkd^dq}{(2\pi)^d}e^{i\mathbf{k}\cdot\mathbf{r}}\langle\theta(\mathbf{k})\,\theta(\mathbf{q})\rangle \\ &= \int\frac{d^dkd^dq}{(2\pi)^d}e^{i\mathbf{k}\cdot\mathbf{r}}\frac{\delta^{(d)}(\mathbf{k}+\mathbf{q})}{\beta\mathcal{J}\epsilon(\mathbf{k})} \\ &= \int\frac{d^dk}{(2\pi)^d}\frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{\beta\mathcal{J}\,|\mathbf{k}|^2}.\end{aligned} \tag{11}$$

This is nothing but a d-dimensional Fourier transform of $|\mathbf{k}|^2$. We now turn to evaluate the correlation function.

- $d=1$ — We may evaluate

$$\langle\theta(\mathbf{r})\,\theta(0)\rangle - \langle\theta(0)\,\theta(0)\rangle \propto \frac{1}{\beta\mathcal{J}}(-|\mathbf{r}|-(-|0|)) \approx \frac{1}{\beta\mathcal{J}}(-r+a)\,, \tag{12}$$

  where $r\equiv|\mathbf{r}|$ and a is the (restored) short-distance cutoff. Notice the fluctuations *diverge* at long-distances.

- $d=3$ — We are familiar with the Fourier transform of $|\mathbf{k}|^2$ in three dimensions, as it is the one obtained from the Coulomb potential $\propto 1/r$,

$$\langle\theta(\mathbf{r})\,\theta(0)\rangle - \langle\theta(0)\,\theta(0)\rangle \propto \frac{1}{\beta\mathcal{J}}\left(\frac{1}{r}-\frac{1}{a}\right), \tag{13}$$

  and the fluctuations become diminishingly smaller at large distances.

- $d=2$ — Let us use a standard trick to calculate

$$\left\langle \theta\left(\mathbf{r}\right)\theta\left(0\right)\right\rangle =\int\frac{dk}{2\pi}\int\frac{d\tilde{k}}{2\pi}\frac{e^{i\tilde{k}r}}{\beta\mathcal{J}\left(k^{2}+\tilde{k}^{2}\right)}\propto\frac{1}{\beta J}\int\frac{dk}{2\pi}\frac{e^{-kr}}{k}\approx\frac{1}{\beta\mathcal{J}}\int_{0}^{r^{-1}}\frac{dk}{2\pi}\frac{1}{k},\tag{14}$$

  and we find

$$\left\langle \theta\left(\mathbf{r}\right)\theta\left(0\right)\right\rangle -\left\langle \theta\left(0\right)\theta\left(0\right)\right\rangle \propto\frac{1}{\beta\mathcal{J}}\log\frac{a}{r},\tag{15}$$

  i.e., we find a logarithmic divergence.

Plugging these results back into the expression for the correlation function we find

$$C\left(\mathbf{r}\right)\propto\begin{cases} e^{-\frac{r}{\#\beta\mathcal{J}}} & d=1\\ \left(\frac{a}{r}\right)^{\frac{\#}{\beta\mathcal{J}}} & d=2\\ e^{\frac{\#}{\beta\mathcal{J}r}} & d=3 \end{cases},\tag{16}$$

which implies that at finite temperature systems in $d=1,2$ are not ordered while those in $d=3$ are. Interestingly, while for $d=1$ the correlation function indeed decays with a scale, for $d=2$ that is not the case. Correlations which decay with a power law are called "algebraic" and can be though of as decaying with a diverging correlation length. Such a phase is sometime called "quasi-long-range ordered" or "critical".

Before concluding this section, let us comment about the classical XY model in $d=2$. We have seen that at low temperature the phase is quasi-long-range-ordered, i.e., exhibits algebraic (power-law) decay of correlations. It is interesting to ask whether at high temperatures a phase transition occurs between this quasi-long-range-order and a fully disordered phase where correlations decay exponentially with some scale. We usually associate a phase transition with symmetry breaking but here neither phases break any symmetry, and so one naively expects no such transition occurs. As it turns out though, there is a phase transition of another kind – a topological phase transition that goes by the name Berezinskii-Kosterlitz-Thouless (BKT). You will study this transition extensively later in the course.

3 Beyond classical XY model

One may think that these results are specific to the classical XY model we have considered, but they are actually quite universal. Any classical system with a continuous symmetry which is spontaneously broken in the ground state will have a massless mode according to Goldstone's theorem. Fluctuations of the order parameter due to this massless mode can then destroy the order, restoring the symmetry, in a similar fashion to what we have showed above – even if the corresponding Hamiltonians are much more complicated. As an example you will study the melting of solids due to phonons (the masless modes of a crystal) in the homework exercise.

Up to now we have only discussed classical problems, however the Mermin-Wagner theorem actually applies to various quantum problems as well. We have seen in the previous tutorial that the partition function of a quantum many-body system can be written in path integral form as

$$\mathcal{Z} = \int \mathcal{D}\left[\bar{\psi}, \psi\right] e^{-\int_0^\beta d\tau \int d^d r \left(\bar{\psi}\partial_\tau \psi + H\left[\bar{\psi},\psi\right] - \mu N\left[\bar{\psi},\psi\right]\right)} . \tag{17}$$

Thinking about the Lagrangian density as an effective classical Hamiltonian density, and about the imaginary time τ direction as another spatial direction, we see that this partition function describes a classical $d+1$ dimensional system which is finite in one direction (the imaginary time direction) and infinite in the d other directions. At zero temperature however, the imaginary time direction become inifinte as well, and so the quantum many-body problem is mapped onto an infinite classical $D = d+1$ dimensional system. This is an example of a quantum-classical mapping which is sometimes possible. Consequently, zero temperature quantum problems in $d = 1$ map to a $D = 2$ classical problem where the Mermin-Wagner theorem applies as we have shown it, namely quantum systems in $d = 1$ are critical. In the same way, a $d = 2$ quantum many-body system at zero temperature maps onto a $D = 3$ classical problem where order can occur. However, at finite temperature the quantum system is more akin to a "thick" 2D classical system, where Mermin-Wagner should apply (for large enough distances).