# The Kosterlitz-Thouless Transition

Conan Weeks*
*Department of Physics and Astronomy, University of British Columbia, Vancouver, BC, V6T 1Z1*
(Dated: November 28, 2005)

A description of phase transitions in two dimensions is presented based on the classical 2D XY-Model. It is shown that the excitation of Goldstone modes in this system rules out spontaneous long-range order. By introducing vortices, it is then shown that a phase transition of a topological nature takes place as the system moves through the critical temperature.

## I. INTRODUCTION

In this paper I shall discuss the fascinating properties of the Kosterlitz-Thouless phase transition in two dimensional systems. In the late 1960's it was known that the appearance of massless goldstone modes (i.e massless excitations that are present whenever the Hamiltonian of a system has a continous global symmetry) precluded the possibility of long-range order at any finite temperature for models with dimension $d \leq 2$. However, numerical studies on 2D spin systems persisted in showing features reminiscent of a phase transition, and thus something other than the traditional order parameter was needed to explain this non-trivial behaviour. Berezinskii (1971) and Kosterlitz and Thouless (1973)[1] were able to show that the phase transition was topological in nature by including vortex excitations in their analysis. At low temperatures it was found that the vortices were bound in pairs, and that a phase transition could take place into a state of unbound vortices as the temperature increased through some critical value. In the following I will give a detailed description of the salient features leading up to their discovery and then briefly introduce more sophisticated methods required to gain a full understanding of the model.

## II. THE 2D XY-MODEL

The Classical 2D XY-Model, which may be considered to be the special case $d = 2$ of the Heisenberg model, is described by the model Hamiltonian

$$H = -J \sum_{\langle i,j \rangle} \mathbf{s}_i \cdot \mathbf{s}_j \tag{1}$$

where J is the ferromagnetic coupling $(J > 0)$ between nearest neighbour spins. The summation is taken over all nearest neighbours on the 2D square lattice, and the spin variables are considered to be classical plane vectors with unit modulus at each lattice site. The spins may therefore be represented as conventional vectors

$$\mathbf{s}_i = S \cos\theta_i \hat{\mathbf{i}} + S \sin\theta_i \hat{\mathbf{j}} \tag{2}$$

$\theta_i$ denoting the angle of the spin on site i with respect to some arbitrary polar direction in the two dimensional vector space, which is a convenient order parameter for the system. Thus

$$\begin{aligned} \mathbf{s}_i \cdot \mathbf{s}_j &= |S|^2 \left(\cos\theta_i \cos\theta_j + \sin\theta_i \sin\theta_j\right) \\ &= \cos\left(\theta_i - \theta_j\right) \end{aligned} \tag{3}$$

and we have

$$H = -J \sum_{\langle i,j \rangle} \cos\left(\theta_i - \theta_j\right) \tag{4}$$

The system is O(2)-Symmetric (i.e a rotation of all the systems spins by the same angle does not change the energy), and since the lowest energy state of the XY-model has all spins pointing in the same direction, the ground state of the XY-model must spontaneously break this symmetry. Now, if we assume that the direction of the spins varies smoothly from site to site, which is reasonable at low temperatures, we can approximate $\cos\left(\theta_i - \theta_j\right)$ by the first two terms in the Taylor expansion, namely, $1 - \frac{1}{2}\left(phi_i - \theta_j\right)^2$, and ignore the rest[4]. This is the spin-wave approximation, named so because of the low energy excitations caused by thermal fluctuations, corresponding to the slow varying rotations of the spins in space - the so called Goldstone modes. It turns out that in any system with a broken continuous symmetry there exist Goldstone modes, which, in any two dimensional system, end up destroying spontaneous order. We note that there are no Goldstone modes when a discrete symmetry is broken, since it is impossible to produce slowly varying rotations from one state to an equivalent one.

So, we are left with the spin-wave approximation of the XY-Model

$$H = -J \sum_{\langle i,j \rangle} \left(1 - \frac{\left(\theta_i - \theta_j\right)^2}{2}\right) \tag{5}$$

which, in the continuum, is equivalent to

$$H = E_0 + \frac{J}{2} \int d\mathbf{r} \left(\nabla\theta\left(\mathbf{r}\right)\right)^2 \tag{6}$$

where $E_0 = -2JN$ is the energy of the completely aligned ground state of N spins and the second term is the good old Mexican-Hat potential (Fig. 1). We shall now find the Magnetization for our system, in order to verify that there is no spontaneous order.

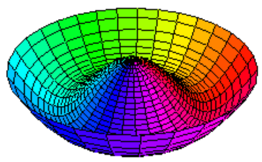

FIG. 1: Goldstone Potential

The thermodynamics of the system will be governed by the partition function[7]

$$Z = \int D\theta \exp\left[-\beta\frac{J}{2}\int d\mathbf{r}\left(\nabla\theta\left(\mathbf{r}\right)\right)^2\right] \tag{7}$$

where $\beta = \frac{1}{k_B T}$ is the Boltzmann factor and we have integrated over all possible configurations

$$\int D\theta = \prod_n \int_{-\pi}^{\pi} d\theta_n \tag{8}$$

We have also discarded the constant term $e^{-\beta E_0}$ since it will cancel out when expectation values are calculated, which, for some observable A, will have the form

$$\langle A\rangle = \frac{1}{Z}\int D\theta A\left(\theta\right)\exp\left[-\beta\frac{J}{2}\int d\mathbf{r}\left(\nabla\theta\left(\mathbf{r}\right)\right)^2\right] \tag{9}$$

Thus, for the x-component of the magnetization we have

$$\begin{aligned} M &= \langle \mathbf{S}_x\rangle \\ &= \frac{1}{Z}\int D\theta \cos\theta\left(\mathbf{r}\right)\exp\left[-\beta\frac{J}{2}\int d\mathbf{r}\left(\nabla\theta\left(\mathbf{r}\right)\right)^2\right] \end{aligned} \tag{10}$$

Now, going over to d dimensions, which will keep the form of our Hamiltonian the same with the exception that the integral over $\mathbf{r}$ will now be d-dimensional, Fourier transforming the phase field

$$\theta\left(\mathbf{r}\right) = \int \frac{d\mathbf{k}}{\left(2\pi\right)^d}\hat{\theta}\left(\mathbf{k}\right)\exp^{-i\mathbf{k}\cdot\mathbf{r}} \tag{11}$$

and substituting back into Eqn. 10, we find the following[4]:

$$\langle \mathbf{S}_x\rangle = exp\left(-\frac{T}{2Jz^{2-d}}S_d\int_{\pi/L}^{\pi/a} dk k^{d-3}\right) \tag{12}$$

For d=2 the integral in Eq. 12 is logarithmically divergent, which is then sufficient to force the magnetization to zero for any non-zero temperature, and thus we cannot have spontaneous order. This is consistent with with our statements above and, more generally, the Mermin-Wagner theorem, which states that there is no spontaneous breaking of a continuous symmetry in systems with short range interactions in dimensions $d \leq 2$.

We will now use the spin-wave expansion at low temperatures and a high temperature expansion of the original XY-Model (Eq.4) to see what changes take place with the spin-spin correlation function

$$\begin{aligned} &\langle \mathbf{s}_0\cdot\mathbf{s}_x\rangle \\ &= \frac{1}{Z}\int D\theta\cos\left(\theta_0-\theta_x\right)\exp\left[-J\beta\sum_{\langle i,j\rangle}\left(\theta_i-\theta_j\right)^2\right] \end{aligned} \tag{13}$$

For low T, we have

$$\langle \mathbf{s}_0\cdot\mathbf{s}_r\rangle \approx \left(\frac{a}{r}\right)^{\frac{1}{2\pi K}} \tag{14}$$

and for high T

$$\langle \mathbf{s}_0\cdot\mathbf{s}_r\rangle \approx \left(\frac{K}{2}\right)^r = \exp^{-r/\xi}, \xi \approx \frac{1}{ln\left(2/K\right)} \tag{15}$$

where $K = J\beta$ and $\xi \approx \frac{1}{ln(2/K)}$[?]

So, we find qualitatively different behaviours for the two cases, namely, an algebraic decay at low temperatures giving rise to a quasi long-range order, as opposed to true long-range order that accompanies a finite magnetization, and an exponential decay at high temperatures corresponding to a short range order. The distinction between the nature of the asymptotic decays allows for a finite temperature phase transition, as promised, and we now need to find some way of accounting for this without symmetry breaking, which will hopefully have something to do with vortices. We first note the failure of the spin-wave approximation to account for the phase transition is rooted in the fact that it only describes the energy cost for small continuous deformations around the ground state. In particular it is ruled out that the field be singular at some isolated point, and so by excluding configurations of spins as seen in Fig. 2, which will lead to such a singularity for a smooth function $\theta\left(\mathbf{x}\right)$, we are excluding from our model all vortex like excitations. For further examples of point singularities see Fig. 3.

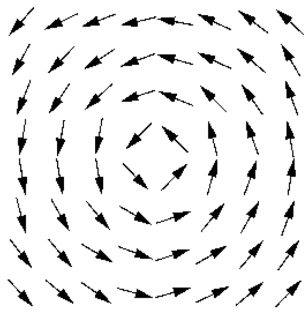

FIG. 2: A vortex

## III. VORTICES

Having arrived at a plausible solution for the mysterious behaviour of the XY-Model, we now want to identify

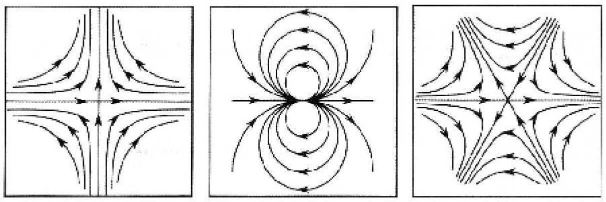

FIG. 3: Point singularities of planar spins in two dimensions with winding numbers (a)-1, (b)2, (c)-2

vortex configurations where the spin field is smooth everywhere except in a small region of space so that we can properly account for their interactions. As noted above, this was exactly the approach taken by Kosterlitz and Thouless who suggested that the disordering is caused by these topological defects, which can not be regarded as simple deformations of the ground state[1]. They have the property that once present in the system, they immediately eliminate the slowly decaying correlations in the spin phase. They also allow for the field to remain smooth everywhere except at isolated fixed points.

Now, since the angle describing the orientation of a spin is defined up to an integer multiple of $2\pi$, it is possible to construct spin configurations in which the traversal of a closed path will see the angle rotate by $\pm 2\pi q$, for some integer q. The integer q measures the 'strength' of the vortex and is referred to as the **topological charge** or **vorticity**, the former being due to the close ties one can make with a 2D coulomb gas. The sign indicates the direction of motion (clockwise rotations being positive). See Fig. 4 for a few possible spin configurations with different values of q. More formally we have the following set of boundary conditions on the circulation of our phase $\theta(\mathbf{r})$[4]: 1) For all close curves encircling the position $\mathbf{r}_0$ of the centre of the vortex

$$\oint_c \nabla\theta(\mathbf{r}) \cdot d\mathbf{l} = 2\pi q \tag{16}$$

2) For all paths that don't encircle the vortex position $\mathbf{r}_0$

$$\oint_c \nabla\theta(\mathbf{r}) \cdot d\mathbf{l} = 0 \tag{17}$$

The contour c may be arbitrary, and both equations are true independent of c, so that q is indeed a characteristic quantity connected with the vortex, and not with the path of integration. Also, if there is more than one vortex inside the contour,the result is the sum of the vorticities, and thus they are additive quantities.

The discrete nature of the topological charge also makes it impossible to find a continuous deformation which returns the state to the uniformly ordered configuration in which the charge is zero. For instance, if we consider a q=1 vortex, like the one in Fig. 2 the spins will rotate through $2\pi$ along every contour enclosing the core, and thus it is impossible to distort the spin configurations into a perfectly aligned state without tampering with them at an arbitrary distance. Any attempt to do so will result in a configuration such as the one in Fig. 5 where boundaries appear at which the spin changes rapidly from up to down, and the only way to achieve the aligned state would then be to have discontinous changes in the direction of the spins. The result is referred to as **Topological Stability**. It is a mathematical concept distinct from physical stability, which depends on free energies of different configurations. However, topological stability usually implies physical stability, as in our case where there is no path to the ground state that is not energetically costly. Our vortex state ends up being stable even though it has a much larger energy than the ferromagnetic ground state.

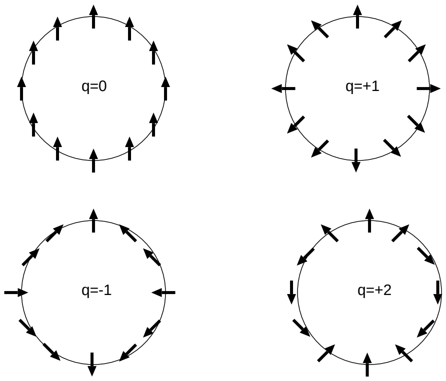


FIG. 4: various vortex configurations

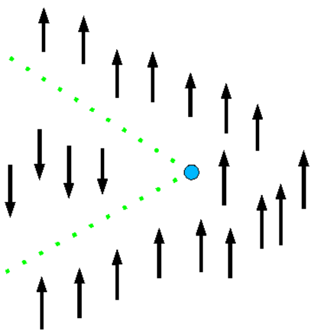

FIG. 5: discontinous spin domains

To estimate the energy (working in the continuum limit), we first calculate the circulation integral along a circle of radius r centred at the position $\mathbf{r}_)$ of the vortex

$$2\pi q = \oint_c \nabla\theta(\mathbf{r}) \cdot d\mathbf{l} = 2\pi |\nabla\theta(\mathbf{r})| \tag{18}$$

where we've used the fact that $\nabla\theta$ , which points in the azimuthal direction, has uniform magnitude because of symmetry. Substituting the result $|\nabla\theta(\mathbf{r})| = q/r$ into

Eqn. 6 we find[4]

$$E_{vor} - E_0 = \frac{J}{2}\int d\mathbf{r}\left[\nabla\theta\left(\mathbf{r}\right)\right]^2$$
$$= \frac{Jq^2}{2}\int_0^{2\pi}\int_a^R rdr\frac{1}{r^2}$$
$$= \pi q^2 J \ln\left(\frac{R}{a}\right) \tag{19}$$

a being the cut off (the core size of the vortex, or, if the model is defined on a lattice,the lattice spacing) and R the system size. We see that the energy diverges logarithmically with R, which prevents the spontaneous formation of defects close to zero temperature. However, by looking at the effect a single vortex will have on the free energy of the system, we can get a rough idea of the temperature at which we might expect to see the proliferation of vortices.

A vortex is characterized by its central co-ordinate, and no matter where it is placed in the system it costs the same amount of energy. So, assuming that the entropy of a vortex is just the logarithm of the number of places we can put it, we should have

$$F_v = \simeq 2\left(\pi J - k_B T\right)\log\frac{R}{a} \tag{20}$$

Therefore, we can reduce the free energy of the system by adding a vortex to a vortex-free system as long as the temperature is above

$$T_c \simeq \pi J/k_B \tag{21}$$

The real system becomes much more complicated when we need to consider more than one vortex, but this argument shows that there may be the possibility of a vortex driven phase transition.

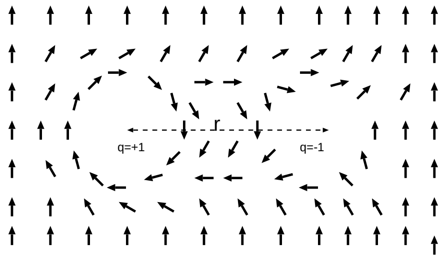


FIG. 6: vortex/anti-vortex pair

Cosider now the combination of a vortex and an anti-vortex, seen in Fig.6. The phase accumulated over a loop enclosing both of them vanishes, and thus we can end up with a far field configuration of the spins that can be distorted continuously to the uniform state. If the vortices are very close, then almost no deformation of the field away from the pair is needed and very little energy is required to create a dipole. So, although isolated vortices can be created and destroyed in a finite system, we shall ignore their effects and treat dipoles as the dominant exciations from the ground state. Quantitatively, we have for the energy of two vortices with respective charges $q_1$ and $q_2$, the expression[3]:

$$E_v \simeq \pi\left(q_1 + q_2\right)^2 \ln\left(R/a\right) + 2\pi q_1 q_2 \ln\left(a/R\right) \tag{22}$$

The first thing to note about this equation is that the lnR divergence is eliminated if $q_1 = -q_2$, or more generally, that there will be no lnR term in a system with multiple vortices if the sum$\sum_\alpha q_\alpha$ of the charges for all vortices is zero. Thus, states satisfying $\sum_\alpha q_\alpha = 0$ will not diverge with R and will be thermally excited for $T > 0$ in agreement with our qualitative argument above. Secondly, the sign of the interaction energy depends on the relative sign of $q_1$ and $q_2$, and since $\ln\left(a/R\right)$ is a monotonically decreasing function of increasing r, the interaction will be repulsive if $q_1$ and $q_2$ have the same sign and repulsive otherwise. To derive these results we shall have to be slightly more rigorous in our approach.

## IV. THE 2D COULOMB GAS

To understand the many vortex interactions, we shall require a model Hamiltonian that actually contains the microscopic characteristics of the system. It turns out we can adapt our continuum model (Eq. 6) to include vortices by splitting the field into two pieces, which is not at all unreasonable, since we know that any vector field can be uniquely decomposed as a sum of irrotational (curl zero) and solenoidal (divergence zero) vector fields. So, doing just that we have

$$\theta\left(r\right) = \bar{\theta}\left(r\right) + \psi\left(r\right) \tag{23}$$

where $\psi\left(r\right)$ is irrotational and can be used to model the spin waves, and $\bar{\theta}\left(r\right)$ is solenoidal and can thus represent the vortices. On the basis of the additivity of the vortices, which was alluded to above, we can introduce the vorticity density in the same way as the notion of charge density is used in electrostatics. Indeed, if we have for the total vorticity within the system

$$q = \sum_i q_i, \tag{24}$$

the vorticity density may be given by

$$\rho\left(r\right) = \sum_i q_i \delta\left(r - r_i\right) \tag{25}$$

where $r_i$ denotes the vortex centre position of the $i$th vortex with vorticity $q_i$, and thus

$$\int d^2 r \rho\left(r\right) = q \tag{26}$$

Actually it turns out that the analogy with electrostatics is total if we let

$$d\bar{\theta}\left(r\right) = \nabla\bar{\theta}\left(r\right) dr \tag{27}$$

play the role of $\mathbf{E}\cdot d\mathbf{l}$ and have $\bar{\theta}$ correspond to the scalar potential $\varphi$. A few lines of algebra will then lead us to a Hamiltonian for a 2D classical neutral plasma containing n positively and n negatively charged particles interacting via a logarithmic coulomb potential[2]

$$H_{2n}\left(r_1, \ldots.., r_{2n}\right) = -\sum_{ij}^{2n} q_i q_j \ln\left(\frac{|r_i - r_j|}{a}\right) + 2n\mu \tag{28}$$

where $\mu$ is the chemical potential, $\sum_i q_i = 0$ and we have only $q_i = \pm q$ charges. This turns out to be just the Hamiltonian we need in order to analyze the critical behaviour of the planar classical XY-Model, and was in fact used by Kosterlitz and Thouless originally to carry out a real space renormalization group analysis of the phase transition. This lead them to the conclusion that the transition could be considered to be of the metal-insulator type. However, we shall not derive the results here and will instead briefly discuss the momentum shell approach of Wilson.

## V. THE RENORMALIZATION GROUP

It turns out that the XY-model can be mapped to its dual Sine-Gordon model, where standard momentum shell RG can then be used to derive the original results of Thouless as well as explain several unresolved questions left from the original real space calculations. The duality comes about from the equivalence between the grand canonical partition function of the Coulomb gas and the canoncial partition function of the 2D Sine-Gordon model. Through feats of analysis that are past my current level of understanding, one can end up with scaling equations for this model, summarily obtain the flow equations, graph the flow diagram (see Fig. 7) and then use the insight gained here to find an equation that determines $T_c$[2]:

$$\frac{\pi J}{k_b T_C} = 1 - 2\pi \exp\left(-\frac{\pi^2 J}{k_b T_C}\right) \tag{29}$$

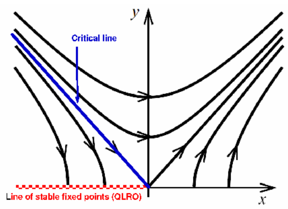


FIG. 7: RG Flow diagram for the Kosterlitz-Thouless phase transition

## VI. SUMMARY

So, in conclusion, we have seen that by including the effects of vortex excitations in the 2D XY-Model, it is possible to account for a phase transition from quasi long range order at low temperatures to disorder at high temperatures, which was not predicted using the simple spin-wave approximation.

---

[*] Electronic Address: `cweeks@physics.ubc.ca`

[1] JM Kosterlitz and DJ Thouless, J. Phys. C: Solid State Phys., Vol.6, 1973.

[2] Z Gulacsi and M Gulacsi, Theory of phase transitions in two-dimensional systems, Advances In Physics, Vol. 47, no. 1, pp.1-89, 1998.

[3] PM Chaitkin and TC Lubensky, Principles of Condensed Matter Physics (Cambridge University Press, Cambridge, England, 1995).

[4] HJ Jensen, The Kosterlitz-Thouless Transition, Department of Mathematics, Imperical College, UK

[5] N Nagaosa, Quantum Field Theory in Condensed Matter Physics (Springer-Verlag Berlin Heidelberg, 1999).

[6] JJ Binney et al., The Theory of Critical Phenomena (Oxford University Press, New York, 1992).

[7] T Korzec, Comparison of exact and numerical results in the XY Model, Dipl. Phys, Humboldt-Universitat zu Berlin, 2003