Introduction to the Berezinskii-Kosterlitz-Thouless Transition

Douglas Packard

May 9, 2013

Abstract

This essay examines the Berezinskii-Kosterlitz-Thousless transition in the two-dimensional XY model. It is an imporant example of phase transitions occuring without an ordered phase or symmetry breaking, and introduced topological ideas into the concerns of condensed matter physicists.

Introduction

Perturbation theory, the technique of choosing an exactly solvable hamiltonian as the jumping off point for a more detailed theory where all additional effects are assumed to be small and approximately calculable with the first couple terms in series expansions, was the heart of many-body physics for such a long time that in some persons' views it was the whole story. The canonical texts (e.g. Abrikosov et. al.) applying (quantum) field theoretic techniques to statistical mechanical and condensed matter problems were largely treatises on perturbation theory, while in high energy physics the particle interpretation made explicit by perturbation theory has guided the field for so long and done so well that many believed field theory itself was nothing more than a means to calculate the collisional cross-sections to be tested at ever-larger accelerator experiments. However, in the case of Landau mean field theory it was realized that perturbation theory could fail in the critical regime where phase transitions are expected to occur, while the particle physicists had spent years trying to grapple with the divergences of perturbations beyond the tree-level. In both cases, the solution lay in a class of techniques now known by the Renormalization Group that involves explicitly admitting the provisional or effective nature of physical theories (i.e. their limited scope) either by ignoring physics beyond a certain energy scale or averaging over short distance scales to derive collective degrees of freedom. The exact understanding of these techniques in the 1970s has changed the way we view physics in such a drastic manner that it is still seeping into the collective understanding of scholars.

However, what still lacked was an understanding of exactly what it was that lay beyond perturbative techniques. In this paper we will examine a model that provided one of the first hints of the richness of physics beyond perturbation theory: the Berezinskii-Kosterlitz-Thouless transition in the two-dimensional XY model. It begins with the discovery of possible field configurations that one could never perturb to from the vacuum and the realization that such states are still physically important, and can in fact dominate observed effects in certain regimes. One finds topologically nontrivial vortex configurations, which are suppressed at low temperatures due to vortex/antivortex binding but become free at high temperatures. Moreover, this is a phase transition that occurs without the spontaneous symmetry breaking familiar from Landau theory.

Two-Dimensional XY Model

The basic model we will be concerned with is that of classical two-component unit vectors (or "spins") interacting in two dimensional space. The general Hamiltonian will be

$$H = -\frac{1}{2}\sum_{r,r'} J(r-r')\mathbf{S}(r)\cdot\mathbf{S}(r') = -\frac{1}{2}\sum_{r,r'} J(r-r')\cos(\theta_r - \theta_{r'})$$

where $\mathbf{S}(r)$ denotes the spin at position r and θ_r the angle of that spin with respect to an arbitrary axis, and we assume $J > 0$. We will be interested in lattice systems with nearest-neighbor interactions, and write the reduced Hamiltonian (hereafter "hamiltonian") as

$$\mathcal{H} \equiv -\beta H = K \sum_{<i,j>} \cos(\theta_i - \theta_j)$$

where $\beta = 1/k_B T$ is the inverse temperature in energy units, $K = \beta J$, and the angle brackets denote a sum over only the nearest-neighbor pairs of lattice points. Next we assume that neighbor spins are close in angle, so we can approximate $\cos(\theta_i - \theta_j) \approx 1 - (\theta_i - \theta_j)^2/2$. Finally we shall take a continuum limit of the lattice, supposing that all other length scales of the model are much greater than the lattice spacing a, hence the angular difference summed over nearest neighbors at each site i becomes a derivative

$$(\theta_i - \theta_{i+\delta x})^2 + (\theta_i - \theta_{i+\delta y})^2 \to a^2(\partial_x\theta_i)^2 + a^2(\partial_y\theta_i)^2 = a^2|\nabla\theta_i|^2$$

and the sum over sites i becomes an integral, giving us an effective hamiltonian

$$\mathcal{H} = E_0 - \frac{K}{2}\int d^2x\,|\nabla\theta(x)|^2$$

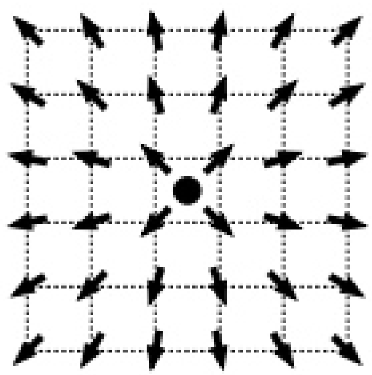
where $E_0 = 2KL^2/a^2$ is the (reduced) energy of the ground state configuration with all spins aligned (assuming a square lattice with area L^2). We ignore the ground state energy, so now this looks like the Gaussian fixed point, but we have a field that we cannot rescale since it is has period 2π. We note that the Mermin-Wagner-Hohenberg theorem states it is impossible for there to be an ordered phase in 2D, and hence no possibility of a phase transition from a disorder to ordered phase accompanied by spontaneous symmetry breaking.

Vortex Configurations

This periodicity allows nontrivial topological configurations: with a standard real-valued field, in order for the field to return to the same value at the end point of a closed path as we move around it, however much the field increases or decreases as we start going around the loop it must reverse course by the time the circuit is finished. For a periodic field however, we can increase or decrease the field by integer multiples of the periodity around the circuit and still come back to the same physical value of the field, which means in this case we can allow θ to change by $2\pi n$ as we go around a loop. Mathematically we can write this constraint as

$$\oint \nabla\theta \cdot d\mathbf{l} = \oint d\theta = 2\pi n.$$

These are the vortex configurations, and an $n = 1$ example is given below. It is clear that it is not possible to continuously deform the vortex field to the constant field configuration of the ground state, hence we expect there to be a conserved topological charge associated with the vortex, which turns out to be n.



What effect do these vortices have on the physics? We take as an ansatz a vortex with charge $n = +1$ located at the origin to have the form $\theta(\mathbf{r}) = \arctan(y/x)$ far from the vortex core (i.e. far enough for our small $\nabla\theta$ approximation to be good). Thus we have $\nabla\theta = (-y/r^2,\, x/r^2)$ and estimate the energy of the single vortex configuration to be

$$U = \frac{J}{2}\int d^2x\,\frac{1}{r^2} = J\pi\int\frac{dr}{r} = J\pi\ln\left(\frac{L}{a}\right)$$

which diverges logarithmically in the thermodynamic limit ($L \to \infty$), hence a lone vortex cannot exist. However, if we pair a $+1$ and a -1 vortex, then far from either of their cores the field will go to a constant value hence the energy should be finite. We can again make an estimate by noting that in the region between the two vortices the angle changes approximately twice as fast, hence

$$U \approx 2J\pi\ln\left(\frac{r}{a}\right)$$

where r is the finite vortex separation, so the vortex and antivortex have a logarithmic attraction (like Coulomb charges in two dimensions).

On the other hand, we can use the single vortex energy to estimate when vortices will become important: there are $(L/a)^2$ ways to place the vortex, hence it has entropy $S = 2k_B \ln(L/a)$, and so free energy

$$F = U - TS = (J\pi - \frac{2}{\beta}) \ln\left(\frac{L}{a}\right).$$

Thus we expect that at low temperatures energy effects will dominate and vortices will be suppressed (the system will behave like a dilute gas of vortex/antivortex pairs) while at higher temperatures the entropic contributions favor the proliferation of vortices. In other words, the pairs will unbind and the system will behave as a vortex plasma (Altland and Simons). We can crudely expect the transition to occur at $T_C = J\pi/2k_B$. This heuristic theory is not reliable for making quantitative predictions, but we shall see it is essentially correct in its qualitative features.

Renormalization Group Analysis

In this section we follow the approach of Altland and Simons in analysing the 2D XY model using RG techniques. We first note that the "distortion" field $\mathbf{u} \equiv \nabla\theta$ is analogous to a velocity field for a fluid, so when there are no vortices the flow should be derivable from a scalar potential: $\mathbf{u} = \mathbf{u}_0 = \nabla\phi$ where ϕ is an unimportant scalar function, so $\nabla \times \mathbf{u}_0 = 0$ and there is no vorticity. We wish to therefore relate the topological charge to the vorticity. This can be done by rewriting the closed path integral constraint as

$$\oint \mathbf{u} \cdot d\mathbf{l} = \int d^2x\, \hat{z} \cdot (\nabla \times \mathbf{u}) = 2\pi n.$$

Hence we can write the vorticity as $\nabla \times \mathbf{u} = 2\pi\hat{z}\sum_j n_j \delta^2(\mathbf{r} - \mathbf{r}_j)$ for a set of vortices of charges n_j at locations $\mathbf{r}_j$. If we then write $\mathbf{u} = \mathbf{u}_0 - \nabla \times (\hat{z}\psi)$ with ψ a scalar field, we find $\nabla \times \mathbf{u} = -\nabla \times \nabla \times (\hat{z}\psi) = \hat{z}\nabla^2\psi$, so

$$\nabla^2\psi = 2\pi \sum_j n_j \delta^2(\mathbf{r} - \mathbf{r}_j).$$

We see then that the field ψ acts as the potential due to charges of strength $2\pi n_j$, and has the solution $\psi(\mathbf{r}) = \sum_j n_j \ln(|\mathbf{r} - \mathbf{r_j}|)$ which is a superpostion of the potentials. We can therefore write any distortion field configuration as

$$\mathbf{u} = \mathbf{u}_0 + \mathbf{u}_1 = \nabla\phi - \nabla \times (\hat{z}\psi)$$

so the hamiltonian $\mathcal{H} = -\frac{K}{2}\int d^2x\, \mathbf{u}^2(\mathbf{r})$ becomes

$$\mathcal{H} = -\frac{K}{2}\int d^2x \left[(\nabla\phi)^2 - 2\nabla\phi \cdot \nabla \times (\hat{z}\psi) + (\nabla \times (\hat{z}\psi))^2\right].$$

We integrate the second term by parts, so assuming no boundary contributions from ϕ it becomes the divergence of a curl and vanishes. Now we simplify the third term by using $\nabla\times(\hat{z}\psi) = -\hat{z}\times\nabla\psi$, so $\nabla\psi$ and $\nabla\times(\hat{z}\psi)$ are perpendicular to each other (and to $\hat{z}$) and both of magnitude $|\nabla\psi|$. Hence $(\nabla\times(\hat{z}\psi))^2 = (\nabla\times(\hat{z}\psi))\cdot(-\hat{z}\times\nabla\psi) = -\hat{z}\cdot(\nabla\psi\times(\nabla\times(\hat{z}\psi))) = (\nabla\psi)^2$, which we then integrate by parts in the hamiltonian and then plug in our above expressions for ψ and $\nabla^2\psi$:

$$\begin{aligned}\mathcal{H}_t &= \tfrac{K}{2}\int d^2x\,\psi\nabla^2\psi \\ &= K\pi\textstyle\sum_{i,j} n_i n_j \ln(|\mathbf{r}_i-\mathbf{r}_j|)\end{aligned}$$

where the subscript on $\mathcal{H}_t$ denotes that it is just the topological term of the hamiltonian. However, this expression includes divergent $i=j$ terms where the logarithmic potential blows up, so it must be regulated by the self (or core) energy of the vortices:

$$\mathcal{H}_t = \sum_i \mathcal{H}^{\rm core}_{n_i} + 2K\pi\sum_{i<j} n_i n_j \ln(|\mathbf{r}_i-\mathbf{r}_j|).$$

So the hamiltonian separates into separate terms for the fields ϕ and ψ, representing the spin-wave and vortex degrees of freedom respectively, hence the partition function factorizes:

$$\mathcal{Z} = \int \mathcal{D}\phi \exp\left[\frac{-K}{2}\int d^2x\,(\nabla\phi)^2\right]\times\sum_{N=0}^{\infty}\frac{1}{(N!)^2}\int\left(\prod_{i=1}^{2N}\frac{d^2x_i}{a^2}\right)e^{\mathcal{H}_t}.$$

We see that in the topological term we are summing over configurations of $2N$ vortices (since we expect them to come in pairs) that act like charges interacting through a Coulomb potential. In fact, if we go back a few steps to the integration by parts, we see that we neglected a boundary term $\oint \psi\nabla\psi d\theta \sim 2\pi\psi(L)\nabla\psi(L)$ as the system size L gets much greater than the positions of any of the vortices. But $\psi = \sum_i n_i \ln(|\mathbf{r}_i-\mathbf{r}_j|) \to \ln L\sum_i n_i$, hence we must have net charge neutrality for this term to vanish: $\sum_i n_i = 0$. It turns out that higher charge vortices are entropically disfavored, so we will work with $n_i = \sigma_i = \pm 1$. This explains the combinatorial factor out front: since the vortices are identical, but must come in pairs that cancel each other out, for any given arrangement $\{\mathbf{r}_i\}$ of the N vortices ($\sigma = +1$), all $N!$ permutations of their labels are equivalent. Similarly for the N antivortices, hence we divide by $(N!)^2$. Furthermore, assuming that the core energy doesn't depend on the sign of the vortex, we write the fugacity as $y_0 = \exp(\mathcal{H}^{\rm core}_{\pm1})$, and hence the full partition function for the topological defects:

$$\mathcal{Z}_t = \sum_{N=0}^{\infty}\frac{y_0^{2N}}{(N!)^2}\int\left(\prod_{i=1}^{2N}\frac{d^2x_i}{a^2}\right)\exp\left[2K\pi\sum_{i<j} n_i n_j \ln(|\mathbf{r}_i-\mathbf{r}_j|)\right]$$

which is the same as that of a neutral Coulomb plasma in two dimensions with unit charges. The spin-wave part of the partition function is Gaussian and has no non-analyticities, hence the phase transition in the 2D XY model must arise from the Coulomb gas term. We can use this to guide our analysis of the phase transition: below T_C we are in an insulating phase where the charge (vortices) are bound together but above T_C the system becomes metallic and the charges are free to move. In other words, the effective interaction of two external charges is screened by the presense of vortices in between them, and above the transition temperature should decay exponentially, allowing the charges to propagate freely.

We calculate the effective interaction perturbatively in the fugacity y_0, only going up to second order where our system has two internal charges located at $\mathbf{s}$ and $\mathbf{s}'$, while the external charges are located at $\mathbf{r}$ and $\mathbf{r}'$. We assume the primed coordinates are negative charges, while the unprimed are positive. Hence (dropping boldface vector notation for convenience)

$$\begin{aligned} e^{\mathcal{H}_{eff}(r-r')} &= \left\langle e^{-2K\pi\ln(|\mathbf{r}-\mathbf{r}'|)}\right\rangle_t \\ e^{\mathcal{H}_{eff}(r-r')+2K\pi\ln(r-r')} &= \frac{1+y_0^2\int d^2s\,d^2s' e^{-2K\pi\ln(s-s')+2K\pi D(r,r',s,s')}+\mathcal{O}(y_0^4)}{1+y_0^2\int d^2s\,d^2s' e^{-2K\pi\ln(s-s')}+\mathcal{O}(y_0^4)} \\ &= 1+y_0^2\int d^2s\,d^2s' e^{-2K\pi\ln(s-s')}\left(e^{2K\pi D(r,r',s,s')}-1\right)+\mathcal{O}(y_0^4) \end{aligned}$$

where $D(r,r',s,s') \equiv \ln(r-s)-\ln(r-s')-\ln(r'-s)+\ln(r'-s')$ and we factored out the $r-r'$ interaction in going from the first to second line, while in going to the third line we rewrote the denominator as $(1+y_0^2[\cdots]+\mathcal{O}(y_0^4))^{-1} = 1-y_0^2[\cdots]+\mathcal{O}(y_0^4)$. We note that the prefactor of the part of the integrand in parentheses acts as a statistical weight that suppresses configurations where the separation $x \equiv s-s'$ is large, we change variables to the relative and center-of-mass coordinates x and $X \equiv (s+s')/2$, or $s = X-x/2$ and $s' = X+x/2$. So we expand in small x, $\ln(r-s) = \ln(r-X) - x\cdot\nabla_X \ln(r-X)/2 + x^2\nabla^2\ln(r-X)/4+\mathcal{O}(x^3)$, and similarly for the other terms in D to find

$$D(r,r',s,s') = -x\cdot\nabla_X\ln(r-X)+x\cdot\nabla_X\ln(r'-X)+\mathcal{O}(x^3)$$

hence

$$e^{2K\pi D(r,r',s,s')}-1 = -2K\pi x\cdot\nabla_X(\ln(r-X)-\ln(r'-X))+4K^2\pi^2[x\cdot\nabla_X(\ln(r-X)-\ln(r'-X))]^2+\mathcal{O}(x^3).$$

To make a long story shorter (see Altland and Simons for further details), we plug this expression back into our expansion in y_0, find the the x linear term vanishes upon integration, then use the Green function property of the Coulomb potential to further simplify things, and finally absorb the short distance cutoff into a rescaling of x. Hence

$$\begin{aligned} e^{\mathcal{H}_{eff}(r-r')} &= e^{-2K\pi\ln(r-r')}\left[1+8\pi^4K^2y_0^2\ln(r-r')\int_1^\infty dx\,x^3e^{-2\pi K\ln x}+\mathcal{O}(y_0^4)\right] \\ &= e^{-2K\pi\ln(r-r')}e^{8\pi^4K^2y_0^2\ln(r-r')\int_1^\infty dx\,x^{3-2\pi K}+\mathcal{O}(y_0^4)}, \end{aligned}$$

so we write the effective hamiltonian $\mathcal{H}_{eff}(r-r') \approx -2K_{eff}\pi\ln(r-r')$ to obtain the effective coupling constant:

$$K_{eff} = K - 4\pi^3K^2y_o^2\int_1^\infty dx\,x^{3-2\pi K}+\mathcal{O}(y_0^4).$$

We see that in the regime where the integrand is finite for $x \to \infty$, the perturbative correction is small, but that the perturbation theory clearly breaks down for $K < K_C = 2/\pi$, which is the same transition point we found using the heuristic free energy argument above (i.e. where the free energy changed sign). Hence we need to be slightly more careful about how we integrate for small K.

This can be done using the renormalization procedure of José et. al. 1977. One integrates only up to a finite distance $x = b$, then re-absorbs this into K, order by order in y_0^2. Hence one has

$$K_{eff}^{-1} = \tilde{K}^{-1} + 4\pi^3 y_0^2 \int_b^\infty dx\, x^{3-2\pi J} + \mathcal{O}(y_0^4)$$

where we used $K_{eff}^{-1} = K^{-1}(1 - K\int_1^\infty[\cdots])^{-1} \approx K^{-1} + \int_1^\infty[\cdots]$ and have defined $\tilde{K}^{-1} \equiv K^{-1} + 4\pi^3 y_0^2 \int_1^b dx\, x^{3-2\pi K} + \mathcal{O}(y_0^4)$. We then rescale $x \to x/b$ to obtain an equation identical to our previous one for K_{eff}^{-1}, but now in terms of shifted and rescaled $K$and y_0:

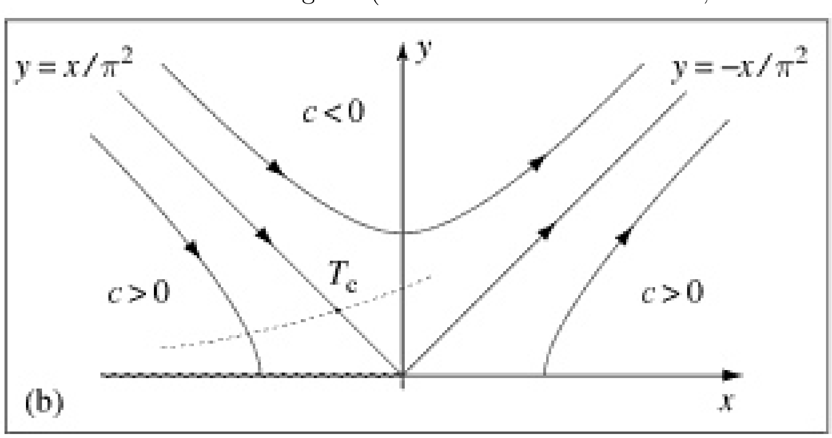
$$K_{eff}^{-1} = \tilde{K}^{-1} + 4\pi^3 \tilde{y}_0^2 \int_1^\infty dx\, x^{3-2\pi\tilde{K}} + \mathcal{O}(y_0^4)$$

where $\tilde{y}_0 \equiv b^{2-\pi K} y_0$. We pick an infinitessimal renormalization $b = e^\ell \approx 1 + \ell$ and hence write the differential RG flow equations

$$\begin{aligned} \frac{dK^{-1}}{d\ell} &= 4\pi^3 y_0^2 \\ \frac{dy_0}{d\ell} &= (2 - \pi K) y_0. \end{aligned}$$

These show that K^{-1} is always increasing with ℓ, while the sign of the derivative of y_0 depends on the coupling K. In particular, we see that at high temperatures (small K), y_0 increases upon renormalization and is therefore a relevant variable, while at low temperatures it is irrelevant, and the transition occurs at $K_C^{-1} = \pi/2$.

Let us look at the flow diagram (source: Altland and Simons; note $x = t$).



For low temperatures and small y (dropping the subscript for convenience), the flow is towards the line of fixed points along $y = 0$, $K^{-1} \leq \pi/2$. This corresponds to the insulating phase, with vortex/antivortex dipoles bound together with some finite radius, hence the fugacity vanishes under renormalization as we coarse-grain the system to larger length scales. The effective interaction is then given by the point at which the flow terminates. Starting from higher temperatures or values of y, however, leads to a flow towards even higher K^{-1} and y, where perturbation theory will break down and we expect free vortices to dominate the physics. The critical point is thus at $(t, y) = (0, 0)$ where $t \equiv K^{-1} - \pi/2$, and the critical trajectory flows into this point. We then reexpress the flow equations near the critical point as

$$\begin{aligned} \frac{dt}{d\ell} &= 4\pi^3 y^2 \\ \frac{dy}{d\ell} &= 4ty/\pi \end{aligned}$$

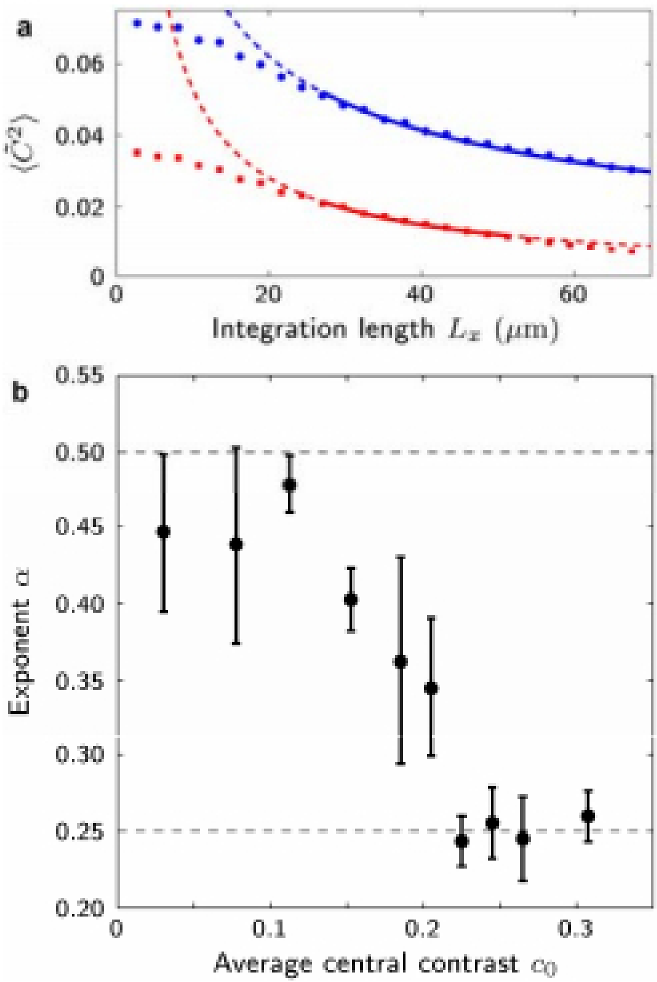
which are manifestly nonlinear recursion relations. To help examine the critical region one can check that the the quantity $c \equiv t^2 - \pi^4 y^2$ is conserved ($dc/d\ell = 0$), hence the flows are characterized by different values of c, each of which is a hyperbola with asymptotes $y = \pm t/\pi^2$, as is apparent from the figure (and the asymptotes themselves correspond to $c = 0$, and are the critical trajectory). So hyperbolae with $c > 0$ correspond to trajectories beneath the critical one, where (starting from low temperature) the flow terminates at ($t < 0, 0$), or starting at $t > 0$ the trajectories flow from $y = 0$ off to infinity. The high temperature case corresponds to the vortex plasma phase since their fugacity is a relevant operator. On the other hand, trajectories with $c < 0$ are above the critical one and correspond to flows from large y at small temperatures, to smaller y as the trajectories flow towards the critical point, but they cross $t = 0$ at $y > 0$ and then head off to infinity. Since the critical trajectory in the low temperature phase is $t = -\pi^2 y$, a nonzero fugacity y_0 thus reduces the critical temperature: $K_C^{-1} = \pi/2 - \pi^2 y_0$.

Before concluding this section, let us return to the topic of order and symmetry breaking in phase transitions. In the high temperature phase of the 2D XY model, we see the expected disorder with exponentially decaying correlations. However, we have seen that there is a phase transition not associated with any symmetry breaking or ordering in the traditional sense: correlations in the low temperature phase are stronger than in the disordered phase, but are still not as strong as the correlations associated with long range order. In fact, it turns out that the low temperature correlations decay with a power law (see Cardy, Altland and Simons). This is known as *quasi-long range order*. Instead of symmetry being the important concept in understanding the degrees of freedom (i.e. seeking an order parameter that quantifies the breaking of symmetry) we are concerned with topological defects in our field configurations and the effects their existence has on the theory. Since the discovery of the BTK transistion, topological ideas have proliferated throughout condensed matter physics much like the vortices in the plasma phase. Through the quantum Hall effect to the today's strange new insulators and superconductors, it is clear that topological thinking is here to stay.

Some Experimental Developments

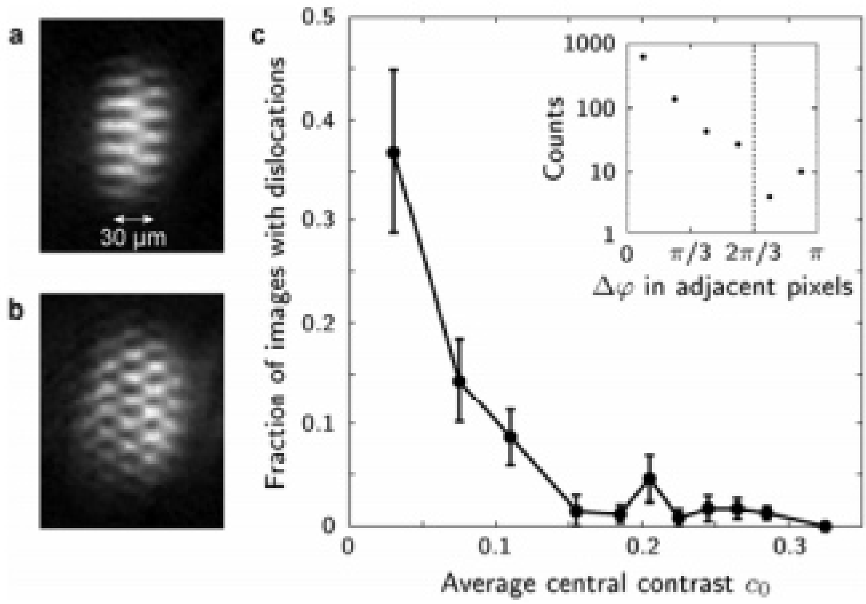
Though originally discovered in the highly idealized 2D XY model, the Berezinskii-Kosterlitz-Thouless transition has proved useful in understanding real experiments due to its relation with the neutral Coulomb gas (as well as to the sine-Gordon model, which we did not mention above). For instance, thin films of superfluids have been shown to have a phase transition with universal properties similar to the BKT transition (see Bishop and Reppy, 1978). Additionally it can be related to a "roughening" transition in crystal surfaces (Cardy). We shall look at an application to trapped atomic gases (Hadzibabic, et. al., 2006).

Here we simply have a gas of quantum degenerate rubidium atoms, trapped in an optical lattice so as to be effectively two-dimensional. These experiments used matter wave interferometry to directly detect free vortices, allowing insight into the vortex binding/unbinding mechanism itself of the BKT transition that had previously been hard to find evidence for. First let us observe the close match between the experimental measurements of the power law correlations and the theoretical predictions:



The top graph shows the agreement of the correlations themselves with the theory, in two different cases, while the bottom graph shows the power law exponents measured.

Next we observe that in these experiments the presense of free vortices in the crossover regime was directly detectable as the vertical dislocations in interference patterns, sometimes showing several free vortices at once!



Perhaps someday we will be able to engineer an experiment that clearly shows all the qualitative effects associated with the vortices in the BKT theory, if only for the beauty of being able to so finely control and understand nature.