Lec. 5.2 B. Notes on the Kosterlitz-Thouless R.G.

Note: Russians and many others call this “Berezinskii-Kosterlitz-Thouless” (BKT) since Berezinskii discovered it independently.

References: My primary reference is Sec. II of the review article by D. R. Nelson in *Phase Transitions and Critical Phenomena*, vol. 7, ed. C. Domb and J. L. Lebowitz (Academic Press, 1983); he goes on to discuss extensions to dynamics, two-dimensional melting, and liquid crystals.

The dimensionless Hamiltonian is given by

\[
\frac{\mathcal{H}}{T} = \int_{\text{except cores}} d^2 r \frac{1}{2} K |\nabla \theta(r)|^2 + N_v E_c
\]  

where \( N_v \) is the number of vortices. The integral is over the whole 2-plane, with the exception of the vortex “cores” which are defined as disks of radius \( a \) around each vortex. The parameter \( a \) is analogous to a lattice constant; not only does it define the vortex core size, but \( a^{-1} \) is also taken to be the wavevector cutoff. (i.e. the same thing as \( \Lambda \) in Lec. 2.7 on the momentum-space R.G.)

As noted in Lec. 5.1, there are two kinds of low-\( T \) excitations:

(i) continuous angle fluctuations, loosely called “spin waves” *

(ii) vortex-antivortex pairs.

Each kind of excitation is governed by one parameter:

(i) \( K = \rho_s / T \) is the dimensionless spin stiffness; in other words, \( K^{-1} \) is a dimensionless temperature variable.

(ii) \( y = e^{-E_c / T} \) is the “vortex fugacity”, where \( E_c \) is the vortex core energy. Thus

\( y \) measures abundance of vortices – \( y = 0 \) means vortices are not allowed at all.

Thus \( y = 0 \) corresponds to the line of critical points in the purely Gaussian theory, done in Lec. 5.1 B. They will be shown to be fixed points of the Kosterlitz-Thouless R.G., so they are critical fixed points. Justified by this fact, we will take \( y \) to be a small parameter: the R.G. is constructed to work in the limit \( y \ll 1 \).

1. Spin map and construction of renormalization group

The renormalization step is to increase \( a \rightarrow ba = (1 + \delta l)a \). What this means practically is that we first draw circles of radius \( ba \) around each vortex. Then if there is a vortex-antivortex pair with \( R_{12} < ba \), we remove the pair and make the angle field smooth in the vicinity. Such a pair is a short-scale fluctuation which is being “integrated out”. It has no net vortex charge: thus, its main effect on larger length scales is to absorb long-wavelength angle gradients at a low energy cost (see (i) below). The last stage of renormalization is to redefine the length scale, \( r' = r / b \), so that finally the new cutoff is the same as the old one, and the new system size is \( L' = L / b \). [This is analogous to the rescaling of the lattice constant in the real-space R.G. (RSRG), and identical to the rescaling done in the momentum-space RG (Lec. 2.7).]

* Do not confuse these with *dynamic* spin waves, as studied in critical dynamics of magnets, e.g. in Chaikin & Lubensky, Sec. 8.3.
Three changes to the Hamiltonian are induced by the spin map.

(i) **Eliminating vortex excitations at small scales**

The close pairs that get eliminated, however, renormalize $K$. Indeed, it was shown in 5.2 A that each pair is analogous to an electric dipole, with $K^{-1}$ being analogous to the dielectric constant of a polarizable medium. Each thermally excited vortex-antivortex pair contributes a term \( \propto R_1^2 \) to the dielectric susceptibility, and hence contributes to the increase of $K^{-1}$. Since $K^{-1}$ is the temperature variable, this means vortices are driving us towards higher $T$, i.e., disorder.

The integral derived in 5.2 A was

\[
K_R^{-1} \approx K^{-1} + 2\pi^2 \left( \frac{y^2}{a^4} \right) \int_a^\infty 2\pi R dR \left[ R^2 e^{-2\pi K \ln(R/a)} \right]
\]

[See Nelson, eq. (2.39)]. Here we used the result for the effective potential \( \exp(-V_{\text{eff}}(R)/T) = \exp(-2\pi K \ln(R/a)) + \text{const} \), and the correct value of “\text{const}” has been absorbed in the prefactors in (2). Also, the factor $y^2 = e^{-2E_c/T}$ is the Boltzmann weight from the cores of the two vortices.

Unfortunately, it was shown in 5.2 A that for sufficiently small $K$ (sufficiently large $T$), the integral (2) diverges. The K.-T. R.G. was invented to fix this. Namely, instead of integrating at once over the entire distribution of pair separations $R$, we integrate step by step; each step only includes the pairs with $a < R_{12} < ba$. You can get the precise result [eq(4a), below] by just plugging into (2) with limits changed to ($a; ab$) and recalling that the integrand is basically constant over that narrow range (since $b - 1 \equiv \delta l \ll 1$.)

**Fig. 1:** steps in K.-T. renormalization

When we consider the statistics of a widely separated pair, as in the upper right corner, this will be controlled by the spin stiffness appropriate to that (nearly macroscopic) length scale. That stiffness will have been renormalized by the close pair fluctuations that occur on short length scales. The K.T. renormalization group asks us to imagine an infinite regression of such pairs-within-pairs. *

* Clearly it won’t be a good approximation when the inner pair has a spacing comparable to the outer one, i.e. when there are four vortices with comparable intervortex distances. The best I can say is that such quartets occur with weight \( \propto y^4 \) and are thus higher order in the small parameter $y$. 
(ii). "Spin-waves" at small scales The modes with wavevectors in \((ba)^{-1}, a^{-1}\) are Gaussians. They are a portion of the \(\mathcal{H}^0\), which are trivial to integrate out, and just give uninteresting constants. [Analogous to \(\mathcal{H}^0\) in momentum-space RG, Lec. 2.7.]

However, there is a less trivial effect. In the "old" system, angle gradients at radius \(r > a\) around a vortex are accounted by the angle gradient energy \(\int \frac{1}{2} \rho_s |\nabla \theta|^2\), but for \(r < a\) they are accounted as part of the core energy. When we increase \(a\), the gradient energy from the annulus \(a < r < ba\) \((1 + \delta l)a\), around each vortex, must be absorbed as part of the new vortex core energy, and accordingly modifies \(y \equiv \exp(-E/T)\)

\[
\delta E_c \approx \pi \rho_s \delta l \Rightarrow \delta y \approx \pi K \delta l \quad y
\]  

(iii). Rescaling and critical behavior

Finally, the rescaling of lengths contributes a factor \(b^2\) to \(y^2\). This just means the density of pairs looks larger by a factor of \(b^2\) in the new system, since we have all the same vortices (apart from a few close pairs, which got integrated out), but the system size is smaller by a factor \(b\). Compare Fig. 1(b) and 1(c). (The same factor of \(b^d\), where \(d = 2\), entered the gradient term too; but it got cancelled exactly by the factor of \(b^{-2}\) from the chain rule in \(\nabla_r \rightarrow \nabla_{r'}\).)

2. The flow equations

The final result is \(K' = K^{-1} + \delta K^{-1}\) and \(y' = y + \delta y\), with

\[
\delta K^{-1} = 4\pi^3 y^2 \delta l \quad (4a) \\
\delta y = (2 - \pi K) y \delta l \quad (4b)
\]

[Reference, Nelson eqs. (2.43)]. Eqs. (4) become continuous (differential) flows in the limit \(\delta l \rightarrow 0\) (i.e. \(b \rightarrow 1\)). The first term ("2") of the prefactor in (4b) comes from contribution (iii) and the second term ("\(\pi K\") comes from contribution (ii). I have left out all terms which are higher order in \(y\).

Observe first that the line \(y = 0\) is indeed a line of fixed points. Also, the factor multiplying \(y\) in (4b) is just the R.G. eigenvalue \(\lambda_y\) at the fixed point. (I mean eigenvalue of the continuous flow.) There is a special "K.T." fixed point at

\[
K^{-1} = K_{KT}^{-1} \equiv \pi/2
\]  

such that \(\lambda_y < 0\) if and only if \(K^{-1} < K_{KT}^{-1}\).

Fig. 2: global flows
You can find that the flow curves are hyperbolas near the K.T. fixed point by linearizing eqs. (4) in its vicinity. A sort of “separatrix” flows into the KT fixed point. The physical curve parametrized by \( (K^{-1}(T), y(T)) = (T/\rho_s, e^{-E_c/T}) \) crosses this “separatrix” at the Kosterlitz-Thouless temperature \( T_{KT} \).

If \( T > T_{KT} \), you flow off to \( (K^{-1}, y) = (\infty, \infty) \) i.e. a disordered state. \((K = 0 \text{ means angle gradients cost nothing, hence their fluctuations are large; also } y = \infty \text{ means } E_c = 0, \text{ so vortices cost nothing and are everywhere.)} \)

\[
\begin{align*}
\eta(T) &
\end{align*}
\]

\[
\begin{align*}
\rho_s(T) &
\end{align*}
\]

If \( T < T_{KT} \), you flow to one of the attracting critical fixed points on the line \( y = 0 \). That means that every temperature below \( T_{KT} \) is a critical temperature. The values \( K^* \) and \( y^* \) at the fixed point which you iterate to, are physically meaningful. The fact that \( y^* = 0 \) means that, at macroscopic scales, there are no vortex pairs in the real system. Also, \( K^* \equiv \rho_R/T \) where \( \rho_R \) is the macroscopic stiffness. It was first measured in a famous experiment by David Bishop (Cornell Ph. D. thesis under J. Reppy, circa 1978). Their realization of K.-T. was a superfluid \(^4\)He film, and \( \rho_R \) is the 2D superfluid density. This was determined by using a torsion oscillator to measure the moment of inertia of a stack of many mylar films, each with a thin helium film on it. Also, the spin-spin correlation function decays at long distances, with the exponent \( \eta \), computed as in Lec. 5.1 B, using \( \rho_R \); thus this exponent varies with temperature, but not quite linearly (as we thought in 5.1 B).

“Universal jump.” Since \( K^*(T_{KT}) = 2/\pi \), it follows that \( \rho_R(T_{KT})/T_{KT} = 2/\pi \) exactly at the Kosterlitz-Thouless transition temperature, a universal ratio (illustrated graphically on see Fig. 4.) Above \( T_{KT} \) the stiffness drops abruptly to zero since a disordered phase has zero stiffness. The universal jump was first measured in the Bishop and Reppy experiment. The correlation exponent \( \eta(T_{KT}) \equiv 1/4 \), right at the \( K - T \) point, is also universal (just insert \( K^*(T_{KT}) \) into the formula from Lec. 5.1 B, \( \eta = 1/(2\pi K) \).

Critical singularities? The singularities at \( T_{KT} \) are essential singularities! The specific heat has a singular part which vanishes as \( \exp(-\text{const}/\sqrt{T - T_{KT}}) \). Likewise, susceptibility and correlation length diverge as \( \exp(+\text{const}/\sqrt{T - T_{KT}}) \). You could check these by integrating the flows (4).

Why it all works: Why is the machinery simpler here than in the momentum-space R.G. (Lec. 2.7)? Well, here the physically interesting fixed point(s) are Gaussian \((y = 0)\); consequently \( y \) is naturally a small parameter. In Lec. 2.7, the interesting Wilson-Fisher fixed point is at \( u \neq 0 \). That involves more complicated (anharmonic) terms; to make \( u \) be a small parameter, we were forced to work in \( d = 4 - \epsilon \) dimensions.