Notes on Classical Topological Phase Transitions with Applications in Soft and Active Matter

Joost de Graaf[1,*]

[1] *Institute for Theoretical Physics, Center for Extreme Matter and Emergent Phenomena, Utrecht University, Princetonplein 5, 3584 CC Utrecht, The Netherlands*

(Dated: March 5, 2026)

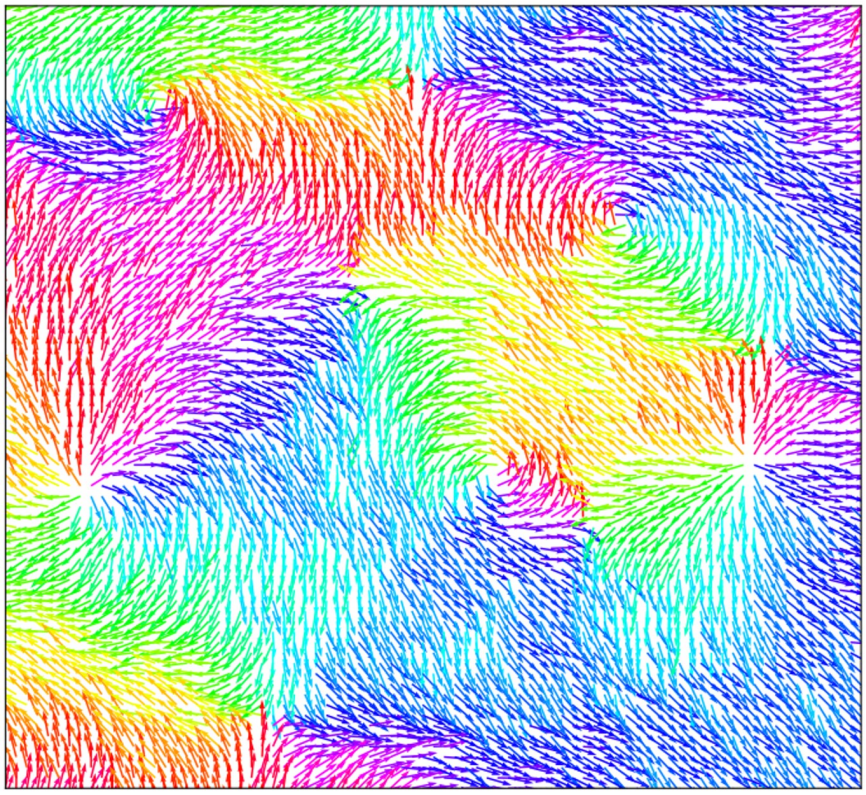

FIG. 1. **The XY model close to the phase transition.** The use of color and arrows indicates the orientation of the field. Note that there are several +1 and -1 charge topological defects visible in the field.

I. INTRODUCTION

In this lecture, we will consider liquid crystals once more. In the first part of the course "Soft and Active Matter Theory," we considered bulk coexistence *via* the order parameter (tensor) S (and $\underline{\boldsymbol{Q}}$). In addition, we have seen how to use Onsager's second-virial theory for isotropic-nematic phase coexistence and how this leads to a self-consistency problem for the orientational density. This may be used to study the coexistence between the isotropic and nematic phases in three-dimensional (3D) lyotropic liquid crystals. This makes it a first-order phase transition.

In two dimensions (2D), the phase transition is not first order, as you have shown in one of the exercises in the first part, due to the symmetry of $\underline{\boldsymbol{Q}}$. In fact, it is a continuous phase transition,

* j.degraaf@uu.nl

which is predicated on local variation of the nematic director and turns out to be a topological phase transition. Here, we will consider this case in more detail by first examining the classical XY model, an extension of the Ising model you are familiar with from your Bachelor's education. In this model, the 'spins' of the Ising model are not only allowed to point up or down, but they have full, in-plane (2π) orientational freedom. We will argue how certain continuous deformations of the orientation of the vector field, save a singular point, give rise to (nearly) zero deformation energy. These points and their properties then lead to a discussion of topology and topological charge. The creation of charge pairs at sufficient temperature will be shown to underly the phase transition in the XY model. The case of liquid crystals in 2D is also discussed as a straightforward extension of the concepts introduced for the XY model. We close by a brief mention of the general relevance of topology to the field of soft and especially active matter.

This lecture aims to make a first entry into the topic of topology from a classical perspective, rather than serve as a full course on the topic. We therefore assume the viewpoint of physicists interested in phase behavior. If you are already familiar with the topic of topology from a more mathematical perspective, you may find some of the discussion lacking rigor. However, we hope that you will nonetheless find the application of topology to phase transitions of interest.

II. THE XY MODEL

To start our discussion, let us introduce the XY model. Like the Ising model, this may be defined on a host of lattices and in a range of dimensions. Here, however, we will be mostly interested in the 2D situation on a square lattice. Consider a system that consists of unit-length spin vectors $\hat{\boldsymbol{s}}_i$, which are free to rotate in the xy-plane. That is, unlike the Ising model for which the spins can point up or down along the y-axis, the spins in this model can be described by an angle $\theta \in (-\pi, \pi]$ with respect to the x-axis. This angle is defined by $\cos\theta_i = \hat{\boldsymbol{s}}_i \cdot \hat{\boldsymbol{x}}$ with $\hat{\boldsymbol{x}}$ the unit vector pointing along the x-axis and '$\cdot$' the inner product. The Hamiltonian for this system can be written as

$$\mathcal{H} = -\frac{J}{2}\sum_i{\sum_j}' \hat{\boldsymbol{s}}_i \cdot \hat{\boldsymbol{s}}_j - \sum_i \boldsymbol{h} \cdot \hat{\boldsymbol{s}}_i = -\frac{J}{2}\sum_i{\sum_j}' \cos(\theta_i - \theta_j) - h\sum_i \cos\theta_i, \tag{1}$$

where J (throughout $J > 0$) is a coupling constant with dimension of energy, $\sum_j'$ indicates a sum over nearest neighbors, and $\boldsymbol{h} \equiv h\hat{\boldsymbol{x}}$ is an external field pointing along the x-axis. The analogy to the Ising model should be obvious.

Exercise 1: Let us make use of our knowledge of statistical mechanics to better understand some of the features of the XY model. For the first few subproblems, we will assume $h = 0$ and that the sites are located on a one-dimensional (1D) chain.

(a) Assume the chain has N sites and is *not* periodic. Sketch a few microstates.

(b) Demonstrate that the expression for the canonical partition sum $Z(N,T)$ for this open-ended 1D chain — N is the total number of sites and T the temperature — can be written as

$$Z(N,T) = 2\pi\left(\int_{-\pi}^{\pi} \mathrm{d}\theta\, \exp\left(\beta J\cos\theta\right)\right)^{N-1} \equiv (2\pi)^N I_0(\beta J)^{N-1}, \tag{2}$$

with $\beta = 1/(k_\mathrm{B}T)$, where k_B is Boltzmann's constant, θ is an auxiliary integration variable, and I_0 is used for notational convenience (it is the modified Bessel function of the first kind).

(c) Compute the free energy per site $f = \beta F/N$ in the thermodynamic limit. Compute the high-temperature limit of f and explain in a *few* words why your results make sense.

(d) Now assume a closed 1D chain, *i.e.*, periodic boundary conditions. Write down the transfer matrix belonging to angles θ and θ' for this problem. How is f (definition in part c) related to the eigenvalues of this matrix? Compute these eigenvalues (using `Mathematica`) and show that this gives the same expression as you obtained in (b) in the thermodynamic limit. Alternatively, skip this part, as you already know that like the Ising model, the XY model in 1D cannot undergo a finite-temperature phase transition.

We now consider the same XY model on a two-dimension (2D) square lattice of N spins. Assume $h \neq 0$ from here on. Define $z_i = \exp(\iota\theta_i)$ with i the index and ι the imaginary unit. Let $z_i = w + \delta z_i$ with $w \equiv \langle z_i \rangle$ the statistical average.

(e) Write down the complex conjugate relations for z_i using the standard $*$ notation. Show that $\cos(\theta_i - \theta_j) = \mathrm{Re}(z_i^* z_j)$, with Re indicating the real part.

(f) Use the result from (e) to expand the Hamiltonian in Eq. (1) up to $\mathcal{O}(\delta z^2)$, *i.e.*, neglecting terms such as $\delta z_i \delta z_j$. Show that the mean-field Hamiltonian is given by:

$$\mathcal{H}^{\mathrm{MF}} = 2JN|w|^2 - 2J\sum_i \mathrm{Re}\left(w^* z_i + w z_i^*\right) - \frac{h}{2}\sum_i \left(z_i^* + z_i\right). \tag{3}$$

(g) Explain in a *few* words why the above free energy should be minimized when w points in the same direction as the external field, *i.e.*, $w \in \mathbb{R}$ and, assuming $h > 0$, $w > 0$. Show that this leads to

$$\mathcal{H}^{\mathrm{MF}}_{\mathrm{min}} = 2JNw^2 - (h + 4Jw)\sum_i \cos\theta_i, \tag{4}$$

The corresponding mean-field free energy per site is given by

$$f^{\mathrm{MF}}_{\mathrm{min}} = 2\beta J w^2 - \log\Big(2\pi I_0\left(\beta(h + 4Jw)\right)\Big). \tag{5}$$

(h) Show that minimizing Eq. (5) leads to $w = I_1\left(\beta(h + 4Jw)\right)/I_0\left(\beta(h + 4Jw)\right)$, with I_1 the derivative with respect to the argument of I_0.

(i) What kind of equation is this for w and what role does w serve in describing the phase transition? Explain in a *few* words.

(j) Based on the above analysis, do you think the 2D XY model has a phase transition? Explain why using only a *few* words.

(k) Make the intuition of subproblem (j) explicit by taking $h \downarrow 0$ and sketching the right-hand side of the equation. Solve for the phase transition temperature T.

III. DEFECTS AND THEIR ENERGY

It turns out that the XY model in 2D has a non-standard phase transition based on the presence of 'defects'. Let us consider these defects next. Similar to the situation of the Ising model, the ground state of the XY model is degenerate. While the former had two possible ground states, all spins up and all spins down, the latter allows for the spins to point in any direction, provided they are all aligned in the same direction ($\boldsymbol{h} = \mathbf{0}$). That is, the configurations in Fig. 2 all have the same (minimum) energy, namely $\mathcal{H} = -2JN$, where we have used that $z = 4$ is the number of neighbors.

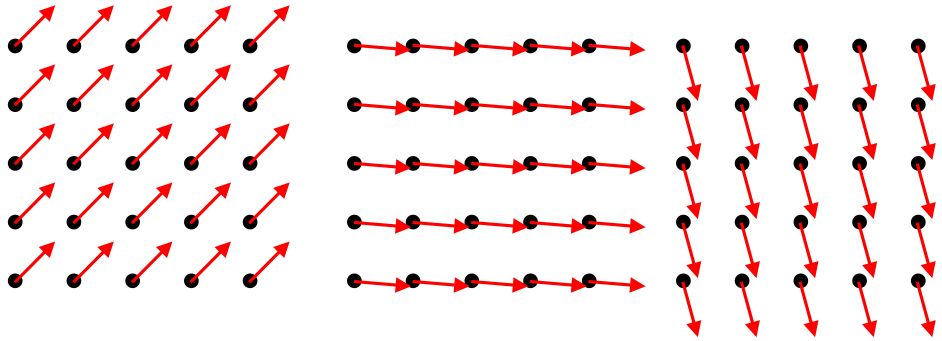

FIG. 2. **The degenerate ground state of the XY model on a square lattice.** Three configurations of the spins $\hat{\boldsymbol{s}}$ (red arrows) that have the same ground-state energy.

We will now look at configurations as a field $\theta(\boldsymbol{x})$ where $x \in \mathbb{Z}^2$ — we can safely assume lattice spacing 1 for now — where θ indicates the angle with x-axis, as before. The directions of the arrows (unit vectors) in the XY model are now given by the vector field $\hat{\boldsymbol{s}}(\boldsymbol{x}) = \big(\cos\theta(\boldsymbol{x}), \sin\theta(\boldsymbol{x})\big) = \hat{\boldsymbol{s}}[\theta]$. Here, the former equality indicates that $\hat{\boldsymbol{s}}$ as a field depends on the position $\boldsymbol{x}$, while it also can be seen to have a functional dependence on $\theta(\boldsymbol{x})$ through the second equality. We will flexibly switch between the two equivalent perspectives.

In this field-based language, the ground state may be described by $\theta(\boldsymbol{x}) = \theta_0$ for some $\theta_0 \in (-\pi, \pi]$. Then it is reasonable to conclude that spatially uniform changes of $\theta(\boldsymbol{x})$ leave the energy unchanged. Think global translations or rotations, where the former are commensurate with the lattice spacing. For example, rotating the $(\boldsymbol{h} = \boldsymbol{0})$ ground state by an arbitrary angle ϕ_0 to $\theta(\boldsymbol{x}) = \theta_0 + \phi_0$ should not change the energy, as Eq. (1) depends on the difference in angle $\cos(\theta_i - \theta_j)$. This intuition also holds for the free energy.

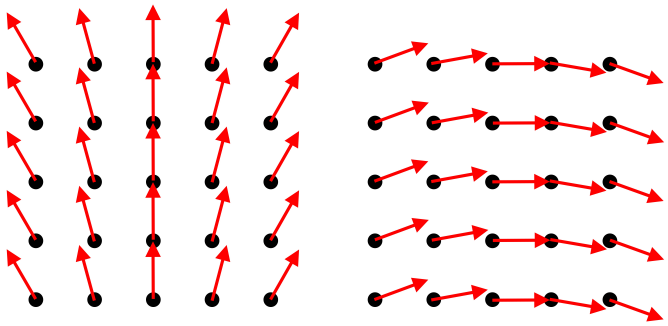

FIG. 3. **Two types of distortion that change the (free) energy of the XY model.** The system is splayed (left) and bent (right). The spins $\hat{\boldsymbol{s}}$ are indicated as red arrows on the 2D square lattice (black).

Non-uniform changes of the field *do* change the (free) energy. Here, you can think of splay and bend, as illustrated in Fig. 3. In the language of fields, we would expect a term to contribute to the free-energy density f that is proportional to $|\nabla\theta|^2$. That is, when the θ_i are sufficiently slowly varying in space that we can sensibly make a continuous approximation and take derivatives, this is the leading-order contribution. We will return to the continuous approximation shortly in an exercise. However, before we do so, note that this energy contribution should look familiar to you, as we recently covered systems with interfaces. In the case of the XY model, this term also serves

to restore the system to a uniform state. As it is the result of deformations, it is often referred to as an elastic contribution. In 3D, there can additionally be a twist contribution to the elastic energy. If you are interested in following up on this, you can read up on the concept of Frank elastic (free) energy, named after Frederick Charles Frank.

Let us take the analysis a bit further before moving onto the exercise. We would expect any free-energy component resulting from elastic deformations to read $f_{\mathrm{el}} = f[\theta(\boldsymbol{x})] - f[\theta(\boldsymbol{x}) = \text{ const.}]$ and this to be analytic in $\nabla\theta(\boldsymbol{x})$. Furthermore, because $\theta(\boldsymbol{x}) = \text{ const.}$ is a minimum of the energy, there should be no linear $\nabla\theta(\boldsymbol{x})$ term present in f_{el}. This leads us to conclude that the simplest form the elastic contribution can take is

$$f_{\mathrm{el}} = \frac{1}{2}\int \mathrm{d}\boldsymbol{x}\, \rho_s |\nabla\theta|^2, \tag{6}$$

where ρ_s is some as of yet unspecified prefactor that sets the system's rigidity and carries the right units. Clearly, in our XY model, this prefactor must be proportional to J.

Exercise 2: In this exercise, we will obtain a field-theoretical form of the XY model.

(a) Taylor expand the zero-field Hamiltonian of Eq. (1) to obtain an expression for the energy of the system in terms of the continuous field $\theta(\boldsymbol{x})$. This should take the form

$$\mathcal{H} \approx E_0 + \frac{J}{2}\int \mathrm{d}\boldsymbol{x}\, |\nabla\theta|^2, \tag{7}$$

by identifying $\theta_i - \theta_j = \partial_x\theta(\boldsymbol{x})$, when i follows j along the x-axis, and similarly for $\partial_y\theta(\boldsymbol{x})$. Here, ∂_x and ∂_y denote the respective partial derivatives along the x- and y-coordinates.

(b) Obtain the criterion $\nabla^2\theta(\boldsymbol{x}) = 0$ for a (local) minimum to $\mathcal{H}$ *via* a functional derivative. Provide a set of solutions to the above condition that globally minimizes the energy. Hint: you do not need to be complete.

The constant solution that you presumably intuited for part (b) of the exercise is not the only solution. We could permit the presence of points, objects with dimension zero (0D), to not satisfy the $\nabla^2\theta(\boldsymbol{x}) = 0$ requirement, such that over (nearly) the entire space the field satisfies the minimization condition. Let us have a look at an archetypal situation where there is a single such point. Imagine the field $\theta(r,\phi) = \phi + \theta_0$ in polar coordinates, *i.e.*, $\boldsymbol{x} = (r,\phi)$, where θ_0 is some constant. Figure 4 shows a few examples. Barring the origin, the associated vector field $\hat{\boldsymbol{s}}(\boldsymbol{x})$ is continuous and smooth. Furthermore, $\nabla\theta(r,\phi) = 1/r$ everywhere away from the origin and $\nabla^2\theta(r,\phi) = -4\pi\delta(r)$. The latter satisfies our $\nabla^2\theta(\boldsymbol{x}) = 0$ save at $\boldsymbol{x} = \boldsymbol{0}$, where the defect is located.

In our continuous variant of the XY model, the troublesome origin can simply be removed from the xy-plane (here we refer to the x- and y-coordinates) by cutting a small disk out of the plane with radius a centered on $\boldsymbol{x} = \boldsymbol{0}$. In that case, on this new coordinate plane (minus origin), the solution is smooth and $\nabla^2\theta(\boldsymbol{x}) = 0$ everywhere. Alternatively, one could have a spatially varying prefactor in front of $\hat{\boldsymbol{s}}$ that goes more rapidly to zero than $1/r$. This factor comes from the fact that the energy integral can be written as

$$\frac{J}{2}\int \mathrm{d}\boldsymbol{x}\, |\nabla\theta|^2 = \frac{J}{2}\int_0^R \mathrm{d}r \int_{-\pi}^{\pi} \mathrm{d}\phi\, \frac{1}{r}, \tag{8}$$

Here, we have considered a finite system[1] with outer radius R to avoid the divergence of the energy as $R \uparrow \infty$. However, the antiderivative of $1/r$ is the natural logarithm, which also diverges for $r \downarrow 0$. Thus, we see that there is an issue with keeping the origin in our coordinate plane.

[1] We are interested in working eventually in the canonical ensemble, which justifies this assumption.

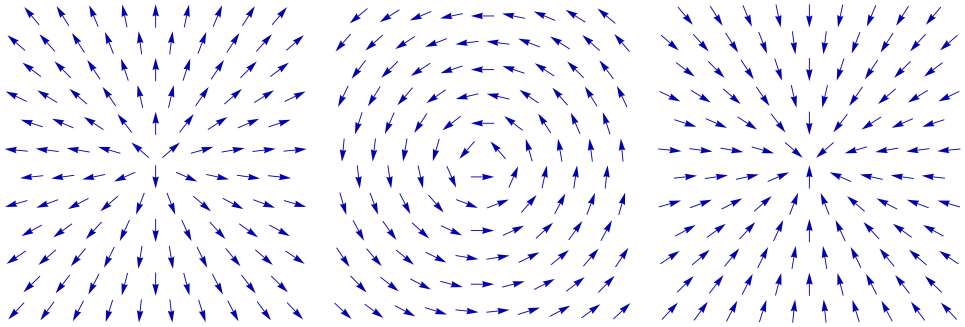

FIG. 4. **The orientation field $\hat{\boldsymbol{s}}(\boldsymbol{x})$ for solutions to $\nabla^2\theta(\boldsymbol{x}) = 0$ containing a single defect.** The blue arrows indicate the direction of the orientation field resulting from $\theta(r,\phi) = \phi + \theta_0$ in polar coordinates (r,ϕ). From left to right the value of $\phi_0 = 0$, $\pi/2$, and π, respectively.

Removing the origin is readily justified physically, as the continuous theory does *not* resolve accurately the orientation of the spins in the XY model below a certain length scale. It will turn out to prove useful to assign a defect a size 'a' and an associated formation energy E_0 — in principle this is dependent on the type of defect that is created. Then, the energy of the system containing a single defect at the origin reads

$$\mathcal{H} = E_0 + \frac{J}{2}\int_a^R \mathrm{d}r \int_{-\pi}^{\pi} \mathrm{d}\phi\, \frac{1}{r} = E_0 + \pi J \log\left(\frac{R}{a}\right), \tag{9}$$

where it should be clear that there is a large energy (and by extension free-energy) cost associated with the introduction of a single defect. This is in spite of the fact that the condition $\nabla^2\theta(\boldsymbol{x}) = 0$ is now satisfied everywhere on the domain. We will return to this in Sections IV and VI.

Clearly, introducing a global bending came at a price and this was to be expected, but this did not reveal itself directly from the shape of the differential equation. However, this is not quite the case, because we have tacitly assumed periodicity. Note that while the orientation field $\hat{\boldsymbol{s}}$ associated with $\theta(\boldsymbol{x})$ is smooth by our choice of polar coordinates, there is a jump in θ of 2π across the half line $y = 0$ and $x < 0$. The placement of this line is arbitrary, but it is the choice we will make throughout. This cut will become important in discussing topology, see Section IV.

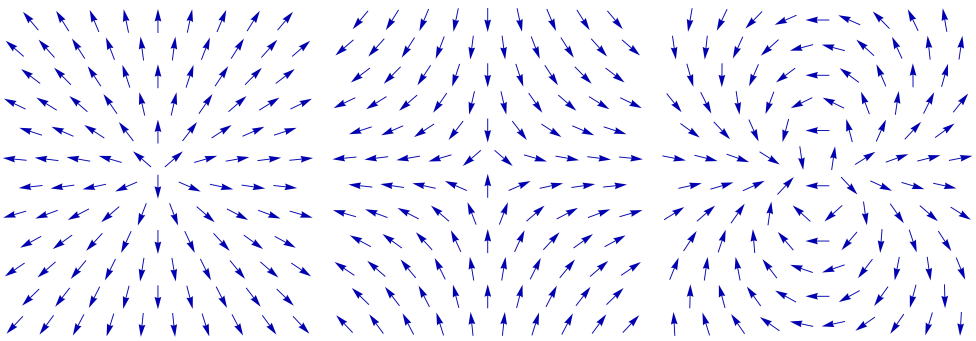

FIG. 5. **Additional solutions to $\nabla^2\theta(\boldsymbol{x}) = 0$ containing a single defect.** The orientation field $\hat{\boldsymbol{s}}(\boldsymbol{x})$ is indicated using blue arrows, when the form $\theta(r,\phi) = k\phi$ is used in polar coordinates (r,ϕ). From left to right the value of $k = 1$, -1, and 2, respectively.

Nothing would have prevented us from using $\theta(r,\phi) = k\phi + \theta_0$ instead, with $k \in \mathbb{Z}\backslash\{0\}$, *i.e.*, an integer other than zero. Figure 5 provides three example vector fields that belong to $k \neq 1$ choices. For the single-defect fields shown in Fig. 5, the above energy calculation simply gives $\mathcal{H} = E_0 + \pi J k^2 \log\left(\frac{R}{a}\right)$, where it is assumed that the E_0 are the same for convenience. This means that the defect energy scales quadratically with the number of times the origin is looped by θ (*e.g.*, at some constant r). Clearly, our intuition should be that such higher order defects are less likely to form 'spontaneously' and in a thermal system we should see less of them.

Before we consider the topology of defects in more detail, let us examine the interaction between defects. It is relatively straightforward to show that the interaction energy between a pair of $\pm k$ defects with a separation vector $\boldsymbol{r}$ is given by

$$\Phi_{\text{pair}} = 2\pi k^2 J \log\left(\frac{|\boldsymbol{r}|}{a}\right), \tag{10}$$

where a is the defect size as before. Determining the force from this pair potential gives an attraction between defects of opposite charge. Like-charge defects repel. In this sense, there is an analogy to electrostatics, which is also borne out by the form of Eq. (10). That is, the pair potential associated with charges in 2D also has a logarithmic character.

Exercise 3: Derive the general expression for the potential energy between two defects of charge k_1 and k_2, respectively, that are positioned $\boldsymbol{r}$ apart.

The way to approach the problem is to position one of the defects at the origin and to say that the other is at $\boldsymbol{r}$ away from the origin, where you align $\boldsymbol{r}$ along the negative x-axis. Next, define $\theta(\boldsymbol{x})$ as the sum of two suitably chosen fields $\theta_1(\boldsymbol{x})$ and $\theta_2(\boldsymbol{x})$. It is not important to write these out, but this can be done by adding two suitably shifted polar representations. Why is it allowed to do add the two contributions like this? Next, use the expression

$$\frac{J}{2}\int_{\Omega} \mathrm{d}\boldsymbol{x}\, |\nabla\theta|^2, \tag{11}$$

to define the energy in the system. Here, Ω is a suitably chosen surface, usually a disk of radius $R \gg |\boldsymbol{r}|$ with two smaller disks cut out of it around the defects. You will need to complement this with two suitably chosen (infinitesimally thin) line cuts that connect these central disks to the outer rim at $|\boldsymbol{x}| = R$. The goal is to cover (nearly) all of the original plane, but leave a closed loop that avoids the defects as the periphery $\partial\Omega$ to Ω. Lastly, make use of integration by parts and completing the square to arrive at the desired expression.

IV. BASIC NOTIONS FROM TOPOLOGY

The value of k is often referred to as the 'charge' of the defect and this is a topological charge. In this section, we will discuss how this comes about and give an impression of the concept of topological protection. We note from Fig. 4 that there are many vector fields that appear to look different but correspond to the same topological charge. Only the phase of the field is different. The value k associated with some vector field that contains a defect can be ascertained directly from the orientation of the vectors $\hat{\boldsymbol{s}}$. That is, we can follow a closed loop around the origin in a counterclockwise fashion, and keep track of the orientation $\hat{\boldsymbol{s}}$ along this path. Then, if this orientation rotates a single time in a counterclockwise fashion, the charge $k = 1$. If it rotates a single time in a clockwise fashion, we have $k = -1$, if it does so twice $k = -2$, and so on. Formulating

this mathematically, we have that

$$k = \frac{1}{2\pi} \oint_c \mathrm{d}\boldsymbol{l} \cdot \nabla\theta(\boldsymbol{x}), \tag{12}$$

where c is the closed path over which is being integrated (for an angle field).

It turns out that it does not matter particularly which closed path is used to obtain the charge, the path itself can be substantially distorted and still give the same value, provided it does not enclose other defects. If it does, then the path measures the total charge that it encloses. The study of this conserved feature is (among other things) the subject of the field of topology. Without doing justice to the complexity of this field, we will now introduce several concepts from it. The intent of this lecture is that a student of this course will have a base grasp of the concept, when they encounter the topic of topology in a research talk. Additionally, they give an interested reader a handle on searching for what else to read.

Crucial to topology are the concepts of 'homotopy' and the 'homotopy group'. Let us start our discussion of these by introducing an order-parameter space $\mathcal{M}$. This is simply a space on which the order parameter field, in our case $\hat{\boldsymbol{s}}(\boldsymbol{x})$ can take its values. For a more 'rigorous' definition, we refer to a basic course on topology. For the XY model, $\mathcal{M}$ is the unit circle S_1. If we were allow the spins to also point out of the plane and assume 3D orientations, then the appropriate $\mathcal{M}$ would be the unit sphere S_2.

Exercise 4: Argue what the order-parameter space $\mathcal{M}$ should be for a 2D nematic liquid crystal. Hint: You do not have to show this, simply argue based on the properties of the liquid crystal. What would the equivalent be for a 3D nematic liquid crystal?

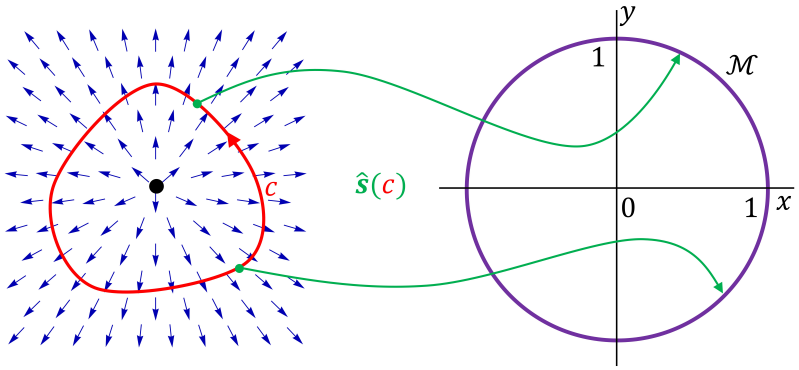

FIG. 6. **Mapping a path c in $\boldsymbol{x}$ onto the order-parameter space $\mathcal{M}$ *via* $\hat{\boldsymbol{s}}$.** (left) The orientation field $\hat{\boldsymbol{s}}(\boldsymbol{x})$ is indicated using blue arrows and the red path c is followed in the counterclockwise direction in the space on which the field lives. The orientation $\hat{\boldsymbol{s}}$ of the field along this taken to the purple unit circle (right). This map can be $\hat{\boldsymbol{s}}[\theta(c)]$, here abbreviated as $\hat{\boldsymbol{s}}(c)$, or some smooth change thereof.

We would say that $\hat{\boldsymbol{s}}$ maps choices of $\theta(\boldsymbol{x})$ onto $\mathcal{M}$. The (co)sines in the definition of $\hat{\boldsymbol{s}}$ make for a smooth function. Pictorially, the action of $\hat{\boldsymbol{s}}$ on a field θ, evaluated along the path c is clarified in Fig. 6. For the path sketched in this figure, the whole space $\mathcal{M}$ is covered once. Hopefully, this makes the discussion in the opening paragraph to this section a bit clearer.

Exercise 5: What would the action of $\hat{\boldsymbol{s}}$ on a path that loops through the ground state be? And what would its action be for a path that does not enclose the defect in Fig. 6? Support your argument using sketches.

In general, a mapping can lead the path to loop $\mathcal{M}$ (here S_1) no time, once, or multiple times (in either direction), depending on whether there are defects enclosed. You should have found in the previous exercise that the ground state is mapped to a point. When it is slightly distorted, for example, by introducing a bit of local splay, any path going through the splayed zone will map out a line segment in $\mathcal{M}$. However, this line segment can be contracted to a point, by removing the local deformation smoothly. Alternatively, you can see this by contracting the entire loop to a point, which gives an analogous single-valued representation in $\mathcal{M}$.

Not so with the +1 topological defect. A loop that circles S_1 once, cannot be continuously contracted to a point, likewise the field cannot be continuously distorted to remove the defect. Analogously, a loop cannot be made to change direction in such a manner. Distortions on top of a topological defect can make the path cover slightly more of S_1, *i.e.*, this mapping will double back on itself. However, these may be smoothly deformed back to a single loop that touches each point of S_1 only once. This loop cannot be made to cover S_1 twice, however, in a continuous fashion. Another way of seeing this is that a defect is associated (in the XY model) with a jump in the value of θ along the negative x-axis. It is these jumps that cause the charge on the defect.

Two paths are now called 'homotopic' if they can be deformed into each other in a continuous manner. Clearly this is true for any path that does not enclose a defect. The presence of defects breaks up the space into a set of maps that are enclosing a (or multiple) defect(s) and ones that do not. All maps that can be identified with each other in a homotopic fashion belong to the same element in a 'homotopy group'. It can be shown that this is a group in the mathematical sense of the word. For our intents and purposes, the charges adhere to a sum rule. That means that a plus and a minus cancel each other out and two +1 defects can be combined into a charge +2 defect.

Exercise 6: Nematic liquid crystals are not like the XY model in one major respect. What are the topological charges that are permitted by the symmetry of these systems? Sketch the lowest charge positive and negative defect.

The fact that one cannot continuously deform one defect into another leads to the idea that topological charge is protected. Let us now develop some physical intuition for this and comment on the validity of thinking of topological protection in an absolute sense in a physical context. Suppose we wish to remove the defect. Then there are several ways to go about this.

The first idea you may have as a physicist is to warm up the system, introduce maximal disorder and then cool it down again into a defect-less state. This would mean changing each spin and from that it is easy to see that the energy cost involved with this process scales as JR^2. Here, R is the size of the system and we note that $JR^2 \gg J\log(R)$ (for $R \gg a$). That is, there is a substantial energy investment required to remove any defect by melting, at least, a much greater one than the cost of keeping the defect around.

Your next idea may be to drag the defect off to the outer edge of the system, where it becomes 'infinity's problem.' Examining, for example, the circular defect in the middle of Fig. 4, we quickly see that this is analogous to cutting open each spin loop. Leaving the defect centered at the origin, this removal strategy is analogous to inserting a wedge of oppositely directed spins into the circle. Figure 7 illustrates this procedure to make the idea clearer. Examining it from the perspective of $\theta(\boldsymbol{x})$ this is the same as backtracking in θ to avoid the jump at the negative x-axis. The energy contribution to doing this can be estimated to scale as JR, since there are strong directional changes of spin along a line. Similar to the melting strategy, $JR \gg J\log(R)$ and this removal approach therefore comes at a high energy cost.

If we do not have (free) energy to spare, then we simply cannot remove the defect and the state is protected. However, for thermal systems the level of protection afforded by topology is not absolute as it would be in the mathematical sense of the word. A final thing to note is that the

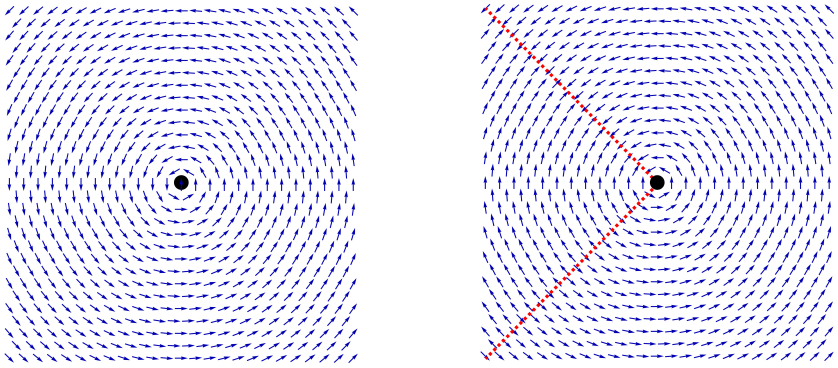

FIG. 7. **Removing the defect via a line.** (left) The orientation field $\hat{\boldsymbol{s}}(\boldsymbol{x})$ is indicated using blue arrows (original, left) and the wedge using a red dashed line (modified, right).

shape of the embedding space, *i.e.*, whether the defects live on a plane or a sphere, controls the total topological charge of the system. If you are interested in learning more about this, you can read up on the Poincaré-Hopf theorem.

V. A PHASE-TRANSITION TOY MODEL

Before we move onto the phase transition in the XY model, let us first consider a simple argument to see that defects can lead to a change in phase as a function of temperature T. In the above discussion, we have mostly focused on deformation energy, but the symmetry and topological arguments can also be employed when we consider the free energy of the system.

Let us assume that the system has a single defect, then the free energy can be written as

$$F = E_{\mathrm{d}} - TS \approx E_0 + \pi J k \log\left(\frac{R}{a}\right) - TS. \tag{13}$$

Here, the entire energy contribution comes from the creation of the defect (with charge k) and the distortion on the field that this induces. The entropy S can be estimated as $k_B \log(R/a)^2$ with k_{B} the Boltzmann constant. This is simply a statement that the defect core can be localized on any one of the $\approx (R/a)^2$ lattice sites. Note that this is, of course, a rather drastic simplification that ignores the edge effects that would come from such a placement in the energy contribution. This constitutes a blunt estimate, and we are not interested in a high level of precision at this moment.

Rewriting the above free energy to $F \approx E_0 + (\pi J - 2k_{\mathrm{B}}T)\log(R/a)$ we realize that $T = \pi J/(2k_{\mathrm{B}})$ sets the free-energy contribution of the defect to a negligible amount E_0. Above this temperature, the contribution of having the defect is negative, meaning that the free energy of the system is lowered by introducing defects. We thus realize that we can go from a defect-free to a defect-rich phase by changing the temperature, which hints at the system undergoing a phase transition. The fact that this temperature is close to the one you obtained from mean-field theory in Exercise 1 is also encouraging[2]. The way in which the system phase transitions is more complex in reality, and we will make this explicit next.

[2] A rightfully critical reader will also note that this is the most sensible combination of parameters that one could put together using dimensional analysis. Thus, it is debatable whether we really learn anything from this argument. Perhaps it is merely a zany way of shoehorning in the entropy concept. The author has included the approach here, as it is found in several textbooks, but leaves it up to the reader to make of it what they will.

VI. DESCRIBING THE PHASE TRANSITION

Returning to the topic of defect interactions, there is a means by which we can 'smoothly' transform away defects or introduce them. This is done by annihilating defects of opposite charge. The opposite procedure, the creation of a pair of defects, is also permitted.

Exercise 7: Graph the vector field caused by two defects that individually have the form $\theta(r,\phi) = k\phi$ with $k = \pm 1$. These are placed at positions $x = \pm\Delta/2$ ($y = 0$) with Δ the separation that can be varied. How does the vector field change upon taking the limit $\Delta \downarrow 0$?

You will have come to the realization that defect unbinding, *i.e.*, the creation of two oppositely charged defects is a way to introduce defects into the system without the messy operations proposed in Section IV. We will now show that defect unbinding is the cause of the phase transition through a (protracted) series of mathematical arguments interspersed with exercises.

Recall that an arbitrary vector field $\boldsymbol{v}(\boldsymbol{x})$ can be decomposed as $\boldsymbol{v}(\boldsymbol{x}) = \boldsymbol{E}(\boldsymbol{x}) + \boldsymbol{H}(\boldsymbol{x})$. Here, the choice of symbols is intentional, and we require that $\nabla \times \boldsymbol{E}(\boldsymbol{x}) = \boldsymbol{0}$ and $\nabla \cdot \boldsymbol{H}(\boldsymbol{x}) = 0$. That is, we employ a Helmholtz decomposition, which you may recall from your electrostatics courses. The $\boldsymbol{E}$ component is irrotational and may be written as $\boldsymbol{E}(\boldsymbol{x}) = -\nabla\Psi(\boldsymbol{x})$ with Ψ a potential. The $\boldsymbol{H}$ field can be written as $\boldsymbol{H}(\boldsymbol{x}) = \nabla \times \boldsymbol{A}(\boldsymbol{x})$ with $\boldsymbol{A}$ a vector potential, for which we impose the gauge $\nabla \cdot \boldsymbol{A}(\boldsymbol{x}) = 0$ throughout to uniquely define this relation. Note that we are working in 2D, hence, we will use the definition of the curl and rotation that extends all vectors by a zero component along the z-axis, *i.e.*, one that is out of the plane.

Exercise 8: We will apply the above considerations to $\nabla\theta(\boldsymbol{x})$ and make use of the intuition that we gained into the problem in the earlier exercises. We start by writing $\theta(\boldsymbol{x}) = \theta_\mathrm{s}(\boldsymbol{x}) + \theta_\mathrm{d}(\boldsymbol{x})$, where the first part comes from small, but removeable, distortions of the ground state, while the second part accounts for the presence of the defects.

(a) Argue that $\nabla\theta_\mathrm{d}(\boldsymbol{x}) = \nabla \times \boldsymbol{A}(\boldsymbol{x})$ based on the analogy to electro/magnetostatics.

To compute $\boldsymbol{A}$ we need to write down a relevant differential equation for the field. Again invoking the analogy, we can use the Maxwell identity $\nabla \times \boldsymbol{H}(\boldsymbol{x}) = \boldsymbol{J}(\boldsymbol{x})$.

(b) Use vector identities to show that $\nabla^2\boldsymbol{A}(\boldsymbol{x}) = -\boldsymbol{J}(\boldsymbol{x})$, which implies that

$$\boldsymbol{A}(\boldsymbol{x}) = \int \mathrm{d}\boldsymbol{x}'\, G(\boldsymbol{x}-\boldsymbol{x}')\boldsymbol{J}(\boldsymbol{x}'), \tag{14}$$

where G is the Green's function belonging to ∇^2, *i.e.*, $\nabla^2 G(\boldsymbol{x}) = -\delta(\boldsymbol{x})$.

(c) Show that $\check{G}(\boldsymbol{k}) = 1/|\boldsymbol{k}|^2$, where $\boldsymbol{k}$ is the Fourier-space vector and therefore

$$G(\boldsymbol{x}) = \frac{1}{(2\pi)^2}\int \mathrm{d}\boldsymbol{k}\, e^{\iota\boldsymbol{k}\cdot\boldsymbol{x}}\check{G}(\boldsymbol{k}) = -\frac{1}{2\pi}\log|\boldsymbol{x}|, \tag{15}$$

where ι is the complex unit as before.

(d) Now determine $\boldsymbol{J}(\boldsymbol{x})$ from the defect positions. Consider a(n arbitrary) loop c an area C that contains N defects at respective positions $\boldsymbol{x}_i$ with i an index and with associated charges k_i. Argue that

$$\sum_{i=1}^{N} 2\pi k_i = \oint_c \mathrm{d}\boldsymbol{l}\cdot\nabla\theta_\mathrm{d}(\boldsymbol{x}) = \int_C \mathrm{d}\boldsymbol{x}\,\nabla\times\nabla\theta_\mathrm{d}(\boldsymbol{x}). \tag{16}$$

(e) Show that the expression from (d) leads to $-\nabla^2 A(\boldsymbol{x}) = 2\pi \sum_{i=1}^{N} k_i \delta(\boldsymbol{x} - \boldsymbol{x}_i)$ and that therefore

$$A(\boldsymbol{x}) = -\sum_{i=1}^{N} k_i \log|\boldsymbol{x} - \boldsymbol{x}_i|, \tag{17}$$

where it is understood that $\boldsymbol{A} = A(x, y)\hat{\boldsymbol{z}}$. The reduced net energy contained in the defects now becomes

$$\beta E_{\mathrm{d}} = \frac{J}{2k_{\mathrm{B}}T} \int \mathrm{d}\boldsymbol{x}\, |\nabla\theta_{\mathrm{d}}(\boldsymbol{x})|^2 = -\frac{\pi J}{k_{\mathrm{B}}T} \sum_{i=1}^{N}\sum_{j=1}^{N} k_i k_j \log|\boldsymbol{x}_i - \boldsymbol{x}_j|, \tag{18}$$

where $\beta = 1/(k_{\mathrm{B}}T)$. This should look familiar, as we expected the net energy of the defects to be given by the sum of the pair potentials we computed in Exercise 7.

(f) Show that Eq. (18) holds using integration by parts and vector identities. What criterion on the k_i follows from the requirement that the boundary term to the integration by parts vanishes? This criterion has a natural physical interpretation in terms of defect unbinding.

(g) Note that the $i = j$ term in the double sum of Eq. (18) causes a divergence. Why is it present? Explain using a *few* words, referencing the approximation made in using $\theta(\boldsymbol{x})$.

As we have seen from the above exercise, we can now write the reduced energy of the entire system as

$$\beta E = -\frac{\beta J}{2} \int \mathrm{d}\boldsymbol{x}\, |\nabla\Psi(\boldsymbol{x})|^2 + \beta E_0 \sum_{i=1}^{N} k_i^2 - 2\pi\beta J \sum_{i=1}^{N}\sum_{i<j}^{N} k_i k_j \log|\boldsymbol{x}_i - \boldsymbol{x}_j|, \tag{19}$$

where the first term accounts for the elastic energy of the removable distortions, the second term for the self-energy (or creation energy) of the defect cores, and the third for the elastic energy of the topological defects. Here, it is assumed that the self energy is equal for all defect types. Now that we have an expression for the energy, we can write down the canonical partition function. It is useful to assume here that the defects come in pairs with ±1 charge. This makes the system satisfy the net-zero charge condition you found in the previous exercise. It is reasonable to assume that only the lowest-order defects contribute significantly to the phase transition, as the higher-order ones have a quadratically increasing elastic energy contribution associated with them.

The grand-canonical partition function for the system, given a chemical potential μ, that adheres to the above assumptions can be written as

$$\begin{aligned}\Xi = \int \mathrm{D}[\Psi]\, \exp\left(\frac{\beta J}{2}\int \mathrm{d}\boldsymbol{x}\, |\nabla\Psi(\boldsymbol{x})|^2\right) \\ \times \sum_{N=0}^{\infty}\left(\frac{1}{N!}\right)^2 \left(\prod_{i=1}^{2N}\int \mathrm{d}\boldsymbol{x}_i\right) \exp\left(-2\beta N\,(E_0 - \mu) + 2\pi\beta J \sum_{i=1}^{2N}\sum_{i<j}^{2N} k_i k_j \log|\boldsymbol{x}_i - \boldsymbol{x}_j|\right),\end{aligned} \tag{20}$$

where the × symbol indicates explicit multiplication and in the first line the integral is over all permissible potentials $\Psi(\boldsymbol{x})$. The first line of Eq. (20) deals with the distortions that cannot be attributed to the presence of defects, while the second line deals with the ones that come from the defects. Note that this is a grand-canonical expression for the number of defects in the system, of which there are $2N$ (N positive and N negative). The positive defects and the negative ones are each indistinguishable. We will henceforth refer to the subpart of Ξ on the first line of Eq. (20) as Z_{s} and the subpart on the second line as Ξ_{d}, so that $\Xi = Z_{\mathrm{s}}\Xi_{\mathrm{d}}$.

Exercise 9: Consider only $Z_{\rm s}$. What do we know about the properties of this part, given the features of the permissible $\Psi(\boldsymbol{x})$? What do you conclude about this part's ability to induce a phase transition? Turning to $\Xi_{\rm d}$, we can shift μ by E_0, so that we have a pure chemical potential dependence. Use this to introduce the fugacity $z = \exp(\beta\mu)$ and rewrite the $\Xi_{\rm d}$.

We can now analyze the defect contribution $\Xi_{\rm d}$ in a controlled approximation that allows us to estimate the transition temperature without invoking renormalization-group arguments. The strategy is to study the stability of the dilute vortex gas by examining the convergence of the grand-canonical partition sum. To do so, we restrict ourselves to defects with charges $k_i = \pm 1$, as argued previously, *i.e.*, higher charges have a quadratically larger elastic energy and are therefore strongly suppressed at low temperatures. We introduce the fugacity $z = \exp(\beta\mu)$ after absorbing the core energy E_0 into the chemical potential. The defect grand-canonical partition function then can be written as

$$\Xi_{\rm d} = \sum_{N=0}^{\infty} \frac{z^{2N}}{(N!)^2} \left(\prod_{i=1}^{2N} \int d\boldsymbol{x}_i \right) \exp\left(2\pi\beta J \sum_{i<j}^{2N} k_i k_j \log|\boldsymbol{x}_i - \boldsymbol{x}_j| \right). \tag{21}$$

By our previous identification, this is also the grand-canonical partition function of a neutral two-dimensional Coulomb gas with associated (appropriate for 2D) logarithmic interactions.

To determine whether the dilute gas is stable, we examine the lowest non-trivial contribution to $\Xi_{\rm d}$, namely the $N = 1$ term, which reads

$$\Xi_{\rm d}^{(1)} = z^2 \int {\rm d}\boldsymbol{x}_1 \int {\rm d}\boldsymbol{x}_2 \, |\boldsymbol{x}_1 - \boldsymbol{x}_2|^{-2\pi\beta J}. \tag{22}$$

We now introduce center-of-mass and relative coordinates $\boldsymbol{R} = (\boldsymbol{x}_1 + \boldsymbol{x}_2)/2$ and $\boldsymbol{r} = \boldsymbol{x}_1 - \boldsymbol{x}_2$, respectively. The Jacobian of this transformation equals unity, and the integral factorizes

$$\Xi_{\rm d}^{(1)} = z^2 \int {\rm d}\boldsymbol{R} \int {\rm d}\boldsymbol{r} \, |\boldsymbol{r}|^{-2\pi\beta J}. \tag{23}$$

The integral over $\boldsymbol{R}$ simply produces the system area A. Introducing a short-distance cutoff a (the defect core size) and an infrared cutoff R (system size), we obtain

$$\Xi_{\rm d}^{(1)} = z^2 A \int_a^R {\rm d}r \, 2\pi r r^{-2\pi\beta J}. \tag{24}$$

The behavior of the radial integral as $R \to \infty$ determines the stability of the dilute vortex gas. We note that if $2\pi\beta J > 2$, the exponent in the integrand satisfies $1 - 2\pi\beta J < -1$ and the integral converges at large R. However, if instead $2\pi\beta J < 2$, the integral diverges as $L^{2-2\pi\beta J}$. Thus, the one-pair contribution per unit area, $\Xi_{\rm d}^{(1)}/A$, remains finite in the thermodynamic limit only if $2\pi\beta J > 2$. This allows us to identify the temperature $k_{\rm B}T_c = \pi J$ as the point at which arbitrarily large vortex-antivortex separations cease to be suppressed[3].

The physical picture is as follows. Below T_c, the dilute expansion is well-defined. Large dipoles are energetically suppressed and vortices remain bound in tightly paired configurations. Above T_c, the partition function density diverges due to contributions from arbitrarily large separations. In this regime, vortex pairs effectively unbind, leading to a finite density of free defects. These free defects destroy the algebraic correlations of the low-temperature phase.

[3] This is quite close to what we already found before using the more jiggery-pokery argument in Section V. Again, it is not surprising to see J and $k_{\rm B}$ playing a major role here. However, it is comforting that the result is now more rigorously derived than we had previously. For a physicist interested in where the phase transition is roughly located, this extra work was clearly not necessary.

It is instructive to compare this result with the behavior of the smooth (defect-free) sector. From the Gaussian integral over $\Psi(\boldsymbol{x})$, one finds $\langle e^{\imath(\theta(\boldsymbol{x})-\theta(0))}\rangle \sim |\boldsymbol{x}|^{-\eta}$ with $\eta = k_{\mathrm{B}}T/(2\pi J)$. In the absence of free vortices, correlations therefore decay algebraically. At the temperature derived above, $\eta_c = 1/2$, which you will recall from our single-defect analysis is where there it becomes favorable to create defects. The defect instability thus coincides with the point at which the algebraic decay exponent reaches a critical value. A more refined treatment, accounting for the renormalization of the stiffness J by vortex fluctuations, yields the universal relation $\eta_c = 1/4$, which characterizes the *Berezinskii-Kosterlitz-Thouless transition*. A full derivation of this universal value requires renormalization-group methods and lies beyond the scope of these notes.

Exercise 10: For enthusiasts of field theory, in this exercise, we investigate the behavior of the smooth (defect-free) sector described by

$$Z_{\mathrm{s}} = \int \mathcal{D}[\Psi] \exp\left(\frac{\beta J}{2}\int \mathrm{d}\boldsymbol{x}|\nabla\Psi(\boldsymbol{x})|^2\right). \tag{25}$$

This is a Gaussian functional integral.

(a) Show that for a Gaussian field one has

$$\left\langle e^{\imath(\Psi(\boldsymbol{x})-\Psi(0))}\right\rangle = \exp\left(-\frac{1}{2}\left\langle(\Psi(\boldsymbol{x})-\Psi(0))^2\right\rangle\right). \tag{26}$$

Hint: Expand the exponential and use Wick's theorem, or recall the corresponding identity for ordinary Gaussian variables.

(b) Using the Fourier representation of Ψ, show that

$$\left\langle\Psi(\boldsymbol{x})\Psi(0)\right\rangle = \frac{k_{\mathrm{B}}T}{J}G(\boldsymbol{x}), \tag{27}$$

where G is the Green's function of the Laplacian,

$$\nabla^2 G(\boldsymbol{x}) = -\delta(\boldsymbol{x}). \tag{28}$$

(c) Using the result $G(\boldsymbol{x}) = -\frac{1}{2\pi}\log|\boldsymbol{x}|$ obtained previously, show that

$$\left\langle(\Psi(\boldsymbol{x})-\Psi(0))^2\right\rangle = \frac{k_{\mathrm{B}}T}{\pi J}\log|\boldsymbol{x}|. \tag{29}$$

(d) Conclude that the spin correlation function behaves as

$$\left\langle e^{\imath(\theta(\boldsymbol{x})-\theta(0))}\right\rangle \sim |\boldsymbol{x}|^{-\eta}, \tag{30}$$

where $\eta = k_{\mathrm{B}}T/(2\pi J)$. What type of order does this correspond to?

Note that this lecture series does not pretend to delve deeply into these aspects of field theory, however, it is instructive and interesting to see that concepts from Statistical Field Theory find application in a Soft Matter context.

We conclude that the phase transition in the two-dimensional XY model can be understood as a thermodynamic instability of the Coulomb gas of defects: above T_c, vortex unbinding proliferates and destroys quasi-long-range order. If you are interested in learning a bit more about the renormalization-group argument, you can turn to the notes "The Kosterlitz-Thouless phase transition: an introduction for the intrepid student" by V. Drouin-Touchette [1].

VII. FINAL CONSIDERATIONS

We hope to have given you a flavor of the role of topology in (2D) phase transitions. If you are interested in the topic, it will turn out that the Kosterlitz-Thouless transition is a much more common feature of 2D systems [2]. This may beg the question, does topology play a role in 3D systems as well? The answer to this is "yes," but in 3D one can also have line-type topological defects [3]. This also leads to connections to knot theory, as these defect lines can form knots. In active system, that is, ones where the liquid crystal can self-propel, the $+1/2$ defect is mobile in 2D. Mobile defects have received interest in the description of the formation of tissues, the spread of cancer, and development of bacterial colonies [4]. More recently, motile 3D defect lines, rings, and knots have seen attention. Lastly, there has recently been a strong interest in shape-from-topology[4]. In these studies, the shape and development of microorganisms is related to the presence of topological defects that must be present on a closed spherical surface [5].

[1] V. Drouin-Touchette, *The Kosterlitz-Thouless phase transition: an introduction for the intrepid student*, arXiv **2207.13748**, 1 (2022).

[2] J. Kosterlitz, *Kosterlitz–Thouless physics: a review of key issues*, Rep. Prog. Phys. **79**, 026001 (2016).

[3] S. Afghah, R. Selinger, and J. Selinger, *Visualising the crossover between 3D and 2D topological defects in nematic liquid crystals*, Liquid Crystals **45**, 2022 (2018).

[4] S. Shankar, A. Souslov, M. Bowick, M. Marchetti, and V. Vitelli, *Topological active matter*, Nat. Rev. Phys. **4**, 380 (2022).

[5] Y. Ravichandran, M. Vogg, K. Kruse, D. Pearce, and A. Roux, *Topology changes of the regenerating Hydra define actin nematic defects as mechanical organizers of morphogenesis*, bioRxiv 2024 (2024).

[4] Not to be confused with geometry from entanglement.