On the theory of superfluidity in two dimensions

Achilles D Speliotopoulos and Harry L Morrison

Department of Physics, University of California, Berkeley, CA 94720, USA

Abstract. The superfluid phase transition of the general vortex gas, in which the circulations may be any non-zero integer, is studied. When the net circulation of the system is not zero the absence of a superfluid phase is shown. When the net circulation of the vortices vanishes, the presence of off-diagonal long range order is demonstrated and the existence of an order parameter is proposed. The transition temperature for the general vortex gas is shown to be the Kosterlitz-Thouless temperature. An upper bound for the average vortex number density is established for the general vortex gas and an exact expression is derived for the Kosterlitz-Thouless ensemble.

1. Introduction

In 1973 Kosterlitz and Thouless published the first of two papers [1, 2] describing a phase transition in two dimensions. Four years later Nelson and Kosterlitz [3] used this theory to predict the celebrated universal jump in the superfluid density of ^{4}He films which was subsequently verified by Bishop and Reppy [4] in their famous torsion pendulum experiment. The cornerstone of this theory is the presence of a *neutral* 'gas' of point vortices in the film which interact through a Coulombic Hamiltonian. A natural analogy is made with a two-dimensional gas of point charges and the phase transition is thought to be due to the binding and unbinding of dipole pairs. The high-temperature phase consists of free vortices each having a circulation or 'charge' $|q| = 1$ while the low-temperature phase consists of oppositely charged vortices bound in dipole pairs. It is thought that the presence of *free* vortices breaks any long range order in the system [5] and, from the Hohenberg-Mermin-Wagner (HMW) theorem [6, 7], we know that the ^{4}He film will behave as a normal fluid. When the temperature is lowered below the transition temperature the gas of vortices condenses into dipole pairs and are no longer free. Long range order is thereby possible. *Renormalization group techniques* are then used to analyse the critical properties of the system and give an operational definition of the superfluid density.

We shall analyse the phase transition in a different, somewhat more traditional way. Unlike Kosterlitz and Thouless we shall *not* begin by differentiating between 'free' and 'bound' vortices but will instead treat them all on an equal footing. We shall apply an infinitesimally small external flow or 'electric' field E_j to the system which may be slowly varying in time. It will serve two purposes. First, it will polarize the system and, as it will be coupled linearly to the polarization vector $P_j = \Sigma \, q_\alpha r_j^\alpha$, the grand partition function becomes a generating functional for correlation functions.

Second, noting that $\tilde{P}_j$ is proportional to the total superfluid momentum density $\tilde{g}_j^{sf}$ in the film, by looking at the superfluid momentum–momentum correlation functions $\langle \tilde{g}_j^{sf} \tilde{g}_k^{sf} \rangle$ we can make contact with the standard two-fluid model in three dimensions. As usual, the normal density will be proportional to the trace of this matrix and will be required to be finite in the static limit $E_j \to 0$. The signature of a superfluid state will be when the off-diagonal terms of the matrix becomes infinite in this limit, signifying off-diagonal long range order. The coefficient of the off-diagonal terms will then serve as our order parameter.

Our approach bares some resemblance to linear response theory. Traditional linear response theory as it is applied to a classical ensemble, however, requires a perturbative expansion in the time derivative of the polarization vector P_j about a small applied field and then letting that field go to zero. For the vortex gas *the polarization vector is a constant of the motion and* $\dot{P}_j$ *vanishes identically.* The perturbation procedure ends before it begins. Fortunately, enough information can be obtained from the grand partition function to demonstrate the presence of long range order as well as the existence of an order parameter without resorting to perturbation theory. Moreover, we find that we need not require that the vortex charges be only ± 1 and will instead begin by considering the more general system in which the charges may be any non-zero integer. *Nor shall we require that the net charge Q of the system to be zero*, although we will consider the cases $Q \neq 0$ and $Q = 0$ separately. Our fundamental premise is that a superfluid state *does* exist in two-dimensional ^{4}He films, and that the Kosterlitz–Thouless–Nelson theory *is* the correct description of it. Our results will, however, require a re-interpretation of the transition itself.

A few words about terminology and notation. By the general Kosterlitz–Thouless (KT) ensemble, or simply the general ensemble, we mean the ensemble in which the vortex circulation may take on any non-zero integer value. The thermodynamic average of an object $\mathcal{O}$ which is calculated using the grand partition function $\mathcal{Z}[|E|]$ for this ensemble will be denoted by $\langle \mathcal{O} \rangle$ and is a function of the applied field E_j. What will be of particular importance is the value it takes in the static limit $E_j \to 0$ which we will denote by the subscript zero: $\langle \mathcal{O} \rangle_0$. We may also calculate the thermodynamic average of $\mathcal{O}$ when $E_j \equiv 0$ identically and this will be denoted by $\langle \mathcal{O} \rangle|_{E_j = 0}$. We shall find that for some $\mathcal{O}$, $\langle \mathcal{O} \rangle|_{E_j = 0} \neq \langle \mathcal{O} \rangle_0$. When we restrict our attention to the KT ensemble, in which all the vortices must have unit charge and the net charge of the system must vanish, we shall denote all thermodynamic quantities calculated using their grand partition function with the index KT. Greek indices will run from 1 to N, the number of vortices in the system, while Latin indices will run from 1 to 2, the dimension of the space. The summation convention will be used throughout this paper when Latin indices are repeated. It will not be used for Greek indices.

The rest of this paper is organized in the following manner. Section 2 provides a brief review of vortex dynamics in two dimensions and gives the motivation for our construction of the grand partition function. In section 3 we shall consider the case in which the net charge does not vanish and show that for this system there is no superfluid phase. The vortices will always behave as a normal fluid. In section 4 we shall consider the case when $Q = 0$ and demonstrate the existence of a superfluid phase transition at the KT temperature T_{KT}. An order parameter for the system is proposed and its relationship with the average density of vortices is derived. All of this analysis is done away from the transition temperature and it is not until section 5 that the behaviour of the system at T_{KT} will be addressed. Concluding remarks may be found in section 6.

2. Background

We begin with a brief description of point vortices in Bose liquids. Let ψ be the microscopic complex scalar field for the ^{4}He atoms. The microscopic current density is

$$j = \frac{\hbar}{2mi}(\psi^\dagger \nabla \psi - \nabla \psi^\dagger \psi) \tag{1}$$

where m is the mass of the ^{4}He atom. We define the superfluid velocity as $v^{\rm sf} = \rho^{-1}j$ [8] where $\rho = |\psi|^2$ is the two-dimensional density of the ^{4}He film. $v^{\rm sf}$ is, as usual, proportional to the gradient of the phase of ψ. As such, for any closed path γ in the film,

$$\oint_\gamma v^{\rm sf} \cdot dl = q\frac{h}{m} \tag{2}$$

where $h = 2\pi\hbar$. Because ψ is a *Bose* field, q may be any integer. When q is non-zero, we say that there is a vortex with *circulation* q somewhere within γ.

At times we shall find it more convenient to work in complex coordinates. We define $z \equiv (x_1 + ix_2)/\sqrt{2}$ with the corresponding definition $v^{\rm sf} \equiv (v_1^{\rm sf} - iv_2^{\rm sf})/\sqrt{2}$. As usual, complex conjugates are denoted by a bar. If we now consider a vortex with circulation q_α located at a position z^α in the film, (2) becomes the contour integral

$$\int_{\gamma_\alpha} v_\alpha^{\rm sf} \, dz = q_\alpha \frac{h}{m} \tag{3}$$

around *any* closed path γ_α in the complex plane encircling z^α. As such, the superfluid velocity for this single vortex is

$$v_\alpha^{\rm sf} = \frac{h}{2\pi im}\frac{q_\alpha}{z - z^\alpha} + v_\alpha^r \tag{4}$$

where v_α^r is any holomorphic, or meromorphic function of z with poles of order greater than one. Since we are primarily interested in the first-order pole of $v^{\rm sf}$, we shall set $v_\alpha^r = 0$. We caution the reader that our definition of the superfluid velocity differs somewhat from that used by Nelson and Kosterlitz [3] but agrees with Minnhagen and Warren [8]. Letting ρ_0 be the spatial average of ρ, we define a total superfluid current density $j^{\rm sf}$ for N vortices located at positions $\{z^\alpha\}$ with circulations $\{q_\alpha\}$ by

$$j^{\rm sf} \equiv \frac{\rho_0 h}{2\pi im}\sum_{\alpha=1}^{N}\frac{q_\alpha}{z - z^\alpha} \tag{5}$$

as well as a superfluid 'momentum' density $g^{\rm sf} \equiv mj^{\rm sf}$.

Looking back at the original proof of the HMW theorem, we see that in addition to Bogoliubov's inequality the conservation equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot j \tag{6}$$

corresponding to the continuous symmetry (in our case a U(1) gauge symmetry) was also used. Because $-\log|z - w|/2\pi$ is the Green function for the Laplace operator in two dimensions, we find that $\nabla \cdot j^{\rm sf} = 0$. As such, the HMW theorem does not take into account vortex excitations in the fluid.

Since $\boldsymbol{\nabla} \cdot \boldsymbol{j}^{\mathrm{sf}} \equiv 0$, while still satisfying (2), the vortices behave as though they are point vortices in an ideal liquid. Treating the vortices as though they were point particles in and of themselves, their Lagrangian is

$$\mathscr{L}_{\mathrm{v}} \equiv -\mathrm{i} \sum_{\alpha}^{N} q_{\alpha} z^{\alpha} \frac{\mathrm{d}\bar{z}^{\alpha}}{\mathrm{d}t} - \mathscr{H} \tag{7}$$

where

$$\mathscr{H} \equiv -e_{\mathrm{KT}} \sum_{\alpha \neq \beta}^{N} q_{\alpha} q_{\beta} U\left(\left|\frac{z^{\alpha} - z^{\beta}}{a}\right|\right) \tag{8}$$

and

$$U(x) = \begin{cases} \log x^2 & \text{for } x \geq 1 \\ 0 & \text{for } x < 1. \end{cases} \tag{9}$$

The reader is referred to either [9] for a standard hydrodynamical derivation of (7) or [10] for a derivation from the microscopic ^{4}He Lagrangian. e_{KT} is the energy scale for the system which we have taken to be $\rho_0 \pi \hbar^2/(2m)$. We caution the reader that our potential U differs from that used by Kosterlitz, Thouless and Nelson by a factor of 2. As usual, a hard core cutoff has been introduced at a distance a to prevent any infrared divergences in the partition function.

Taking our generalized coordinate to be $\bar{z}^{\alpha}$, its canonical momenta is then $-\mathrm{i}q_{\alpha} z^{\alpha}$ and we see immediately that $\mathscr{L}_{\mathrm{v}}$ is linear in the momenta. Consequently, the vortex Hamiltonian $\mathscr{H}_{\mathrm{v}}$ obtained from $\mathscr{L}_{\mathrm{v}}$ is simply $\mathscr{H}$. It contains no 'kinetic' piece and is purely 'potential'.

The symmetries of $\mathscr{L}_{\mathrm{v}}$ are well known and listed in table 1 (see either [11] or [12] for a somewhat different approach). Corresponding to rotational invariance there is the total angular momentum

$$I \equiv \rho_0 h \sum_{\alpha}^{N} q_{\alpha} |z^{\alpha}|^2 \tag{10}$$

while invariance under translations gives the total linear momentum

$$P \equiv \rho_0 h \sum_{\alpha}^{N} q_{\alpha} z^{\alpha} \qquad \bar{P} \equiv \rho_0 h \sum_{\alpha}^{N} q_{\alpha} \bar{z}^{\alpha}. \tag{11}$$

P has the *form* of a polarization vector which justifies calling it such. It is also related to the total superfluid momentum density g^{sf} in the following way. Let $\mathscr{D}$ be a disk in

Table 1. Table of the symmetries of the vortex Lagrangian and its conserved charges.

Symmetry	Conserved quantity	Observable	Lagrange multiplier	Physical interpretation
Time translation invariance	$\mathscr{H}$	Energy	$\beta = 1/T$	Temperature
Space translation invariance	$P, \bar{P}$	Superfluid current/ polarization vector	$\bar{E}, E$	External flow/ electric field
Rotational invariance	I	Angular momentum	Q	Total vorticity/ charge

the complex plane such that $z^\alpha \in \mathcal{D}$ for all α. Then using the identity (see the appendix)

$$\bar{z}^\alpha = \frac{1}{2\pi i} \int_{\mathcal{D}} \frac{dz \wedge d\bar{z}}{z - z^\alpha} \tag{12}$$

which is independent of the size of $\mathcal{D}$, we find that in rectangular coordinates

$$P_j = -\varepsilon_{jk} \int_C g_k^{\text{sf}} d^2 r \tag{13}$$

where ε_{jk} is the totally anitsymmetric pseudotensor $(\varepsilon_{jk} = -\varepsilon_{kj}, \varepsilon_{12} = 1)$.

The grand partition function for the ensemble is then

$$\mathcal{Z}[E, \bar{E}] \equiv \sum_{\text{config}} g(N, q)\lambda(N, q)Z_N[E, \bar{E}] \tag{14}$$

where

$$Z_N[E, \bar{E}] \equiv \prod_\alpha^N \int_C \frac{i\, dz^\alpha \wedge d\bar{z}^\alpha}{a^2} \exp\left(-\beta\left[\mathcal{H} - (\bar{E}P + E\bar{P}) + \frac{\rho_0 h}{m} QI\right]\right) \tag{15}$$

$(\bar{E}P + E\bar{P} = E_j P_j)$, is the N-vortex partition function and the sum is over all configurations of the system. $g(N, q)$ is the multiplicity factor due to the number of identical vortices in each configuration while $\lambda(N, q)$ represents the total fugacity of the system. Their exact forms are quite complicated, but fortunately are not needed for our purposes. Q and E_j are the Lagrange multipliers corresponding to the conserved charges I and P_j respectively. Physically, Q is interpreted as the total vorticity (angular velocity) of the system. Since even in the absence of a net external rotational flow in the system there still may be a net vorticity due to that of the vortices themselves, $Q = \sum q_\alpha$. This is the case we shall always consider. E_j is identified as the components of a net external non-rotational flow or 'electric' field which may be slowly varying in time. In complex coordinates $E \equiv (E_1 + i E_2)/\sqrt{2}$.

First, we note that the grand partition function is a function of the magnitude of E_j only and not its direction: $\mathcal{Z}[E, \bar{E}] = \mathcal{Z}[|E|]$ where $|E|^2 \equiv \bar{E}E = E_j E_j / 2$. This is due to the rotational invariance of $\mathcal{H}$ and I. Second, in the absence of any external field whatsoever, $\langle P_j \rangle|_{E_j = 0} = 0$ for all values of Q; once again due to the rotational symmetry of $\mathcal{H}$ and I. Such a state, however, is never realized experimentally and, when $Q = 0$, is, as we shall see, extremely unstable. Since in reality there is always some external current flow in the system, we will introduce an external flow to the system and look once again at the behaviour of $\langle P_j \rangle$ when $E_j \to 0$. The system will behave very differently depending upon whether or not Q vanishes and we will treat the two cases separately.

3. The case $Q \neq 0$

Turning our attention to the N-vortex partition function, we complete the square by letting

$$w^\alpha = z^\alpha - \frac{mE}{\rho_0 h Q}. \tag{16}$$

Because $\mathcal{H}$ is translationally invariant,

$$\mathcal{Z}[|E|] = \exp\left(\beta \frac{m}{2} E_j E_j\right) \mathcal{Z}[0] \tag{17}$$

and it is now straightforward to calculate the average polarization vector

$$\langle P_j \rangle = m E_j. \tag{18}$$

The fluctuation $\tilde{P}_j \equiv P_j - \langle P_j \rangle$ about this average flow is

$$\langle \tilde{P}_j \tilde{P}_k \rangle = \frac{m}{\beta} \delta_{jk}. \tag{19}$$

From (13), and using translational invariance in the $E_j \to 0$ limit,

$$\langle \tilde{P}_j \tilde{P}_k \rangle_0 = \Omega \varepsilon_{jl} \varepsilon_{km} \int_\Omega \langle \tilde{g}_l^{\mathrm{sf}}(\boldsymbol{r}) \tilde{g}_m^{\mathrm{sf}}(0) \rangle_0 \, \mathrm{d}^2 \boldsymbol{r} \tag{20}$$

where Ω is the total surface area of the film. As usual we define the normal fluid density as

$$\rho_n \equiv \frac{\beta}{2m} \int_\Omega \langle \tilde{g}_j^{\mathrm{sf}}(\boldsymbol{r}) \tilde{g}_j^{\mathrm{sf}}(0) \rangle_0 \, \mathrm{d}^2 \boldsymbol{r}. \tag{21}$$

Using (19) and (20) to calculate ρ_n, we immediately see that the system will always behave as a normal fluid. *There is no superfluid phase when $Q \neq 0$.* Note also that for this case $\langle P_j \rangle|_{E_j = 0} = \langle P_j \rangle_0$ and, not surprisingly, rotational symmetry is restored when $E_j \to 0$.

4. The case $Q = 0$

We now sum only over those configurations of the system for which the net circulation vanishes. The term proportional to I is no longer present in the Boltzmann factor so that neither $\langle P_j \rangle$ nor $\langle \tilde{P}_j \tilde{P}_k \rangle$ is trivial to calculate. We can, however, get some notion of their behaviour in the static limit by looking at the properties of $\langle \delta \mathcal{H} \rangle = -\langle E_j P_j \rangle$. We begin by establishing a very important inequality.

Returning once again to the grand partition function, we scale the coordinates of the vortices by letting $w^\alpha = \bar{E} z^\alpha$ so that $\mathrm{i}\, \mathrm{d}w^\alpha \wedge \mathrm{d}\bar{w}^\alpha = |E|^2 \mathrm{i}\, \mathrm{d}z^\alpha \wedge \mathrm{d}\bar{z}^\alpha$. Then

$$\mathcal{Z}[|E|] = \sum_{\mathrm{config}} \mathrm{g}(N, q) \lambda(N, q) \exp\left(-2 \left[N - \frac{\beta}{\beta_{\mathrm{KT}}} D \right] \log|E| \right) Z_N[1] \tag{22}$$

where $D \equiv \sum q_\alpha^2$ and $\beta_{\mathrm{KT}} \equiv 1/e_{\mathrm{KT}}$ is the KT temperature. In obtaining this expression we have made use of the identity

$$0 = \left(\sum_\alpha^N q_\alpha \right)^2 = D + \sum_{\alpha \neq \beta}^N q_\alpha q_\beta. \tag{23}$$

Then

$$\beta \langle \delta \mathcal{H} \rangle = -\left(E \frac{\partial}{\partial E} + \bar{E} \frac{\partial}{\partial \bar{E}} \right) \log \mathcal{Z} = 2 \left(\langle N \rangle - \frac{\beta}{\beta_{\mathrm{KT}}} \langle D \rangle \right). \tag{24}$$

Since $|q_\alpha| \geq 1$ for all α, $D \geq N$ and

$$\beta \langle \delta \mathcal{H} \rangle \leq 2 \left(1 - \frac{\beta}{\beta_{\mathrm{KT}}} \right) \langle N \rangle. \tag{25}$$

Using this inequality and the observation that $\langle N \rangle \geq 0$ it is straightforward to show that $\langle \delta \mathcal{H} \rangle$ has the following properties:

1. $\langle \delta \mathcal{H} \rangle |_{E_j = 0} = 0$,
2. $\langle \delta \mathcal{H} \rangle \leq 0$ for all $\beta \geq \beta_{KT}$ ($T \leq T_{KT}$), $E_j \neq 0$,
3. if $\langle \delta \mathcal{H} \rangle = 0$ for some $\beta > \beta_{KT}$ ($T < T_{KT}$), $E_j \neq 0$, then $\langle N \rangle = 0$ there,
4. if $\langle \delta \mathcal{H} \rangle > 0$ for some $\beta < \beta_{KT}$ ($T > T_{KT}$), $E_j \neq 0$, then $\langle N \rangle > 0$ there,

where for completeness we have included the result obtained in section 2.

In the absence of any vortices whatsoever, the HMW theorem prevents a phase transformation from taking place and the ^{4}He film will behave as a normal fluid. Since *the low-temperature phase is a superfluid, we will assume that* $\langle N \rangle \neq 0$ when $\beta > \beta_{KT}$. From properties 2 and 3 we find that

$$\langle \delta \mathcal{H} \rangle_0 \equiv \lim_{E_j \to 0} \langle \delta \mathcal{H} \rangle < 0 \tag{26}$$

for all $\beta > \beta_{KT}$. *A priori* there is no reason why this limit will exist. If, however, $\langle \delta \mathcal{H} \rangle \to -\infty$, as $E_j \to 0$, we would then have the unphysical result that an infinitesimally small external field causes an infinitely large shift in the average energy of the system. On physical grounds we conclude that $\langle \delta \mathcal{H} \rangle_0$ must be finite. Moreover, we will find that in order to define a normal fluid density we must also require that $\langle \delta \mathcal{H} \rangle$ be expandable in a Taylor series about $E_j = 0$.

We now see that due to the presence of vortices in the low-temperature phase, $\langle \delta \mathcal{H} \rangle |_{E_j = 0} \neq \langle \delta \mathcal{H} \rangle_0$. Even when the applied field is turned off rotational symmetry remains broken so that the low-temperature phase is in a state of broken symmetry. Returning for a moment to the definition of the superfluid velocity, we let ϕ be the phase of ψ. From the definition of v^{sf} and (4),

$$\phi = -i \log \left[\prod_{\alpha=1}^{N} \left(\frac{z - z^\alpha}{\bar{z} - \bar{z}^\alpha} \right)^{q_\alpha} \right] \tag{27}$$

where we have once again set all v_α^r to zero. Now perform a global gauge transformation $\psi \to \exp(iN\chi)\psi$. Then $\phi \to \phi + N\chi$, or, equivalently, $z^\alpha \to \exp(i\chi/2q_\alpha)z^\alpha$. A global gauge transformation of the Bose field is the same as a uniform rotation of the vortices. Breaking rotational invariance is equivalent to breaking the U(1) gauge symmetry so that like superfluidity in three dimensions, superfluidity in two is also characterized by the breaking of a U(1) gauge symmetry. Furthermore, because $\langle \mathcal{H} \rangle_0 + \langle \delta \mathcal{H} \rangle_0$ is simply the average energy of the broken symmetry state, the broken symmetry state has less energy than the symmetric one $\langle \mathcal{H} \rangle |_{E_j = 0}$ and is the one favoured energetically.

For convenience we define $R = -\beta \langle \delta \mathcal{H} \rangle$. It is a function of the external field E_j so that by expanding R in a Taylor series,

$$R = R_0 + \frac{1}{2!} \frac{\partial^2 R}{\partial E_k \partial E_k} \bigg|_0 E_j E_j + \text{HOT}. \tag{28}$$

There are no terms in the expansion which contain odd powers of E_j since $\mathcal{H}$ is a function of $|E|$ only. The coefficient of each term in the expansion, although independent of E_j, is still a function of β and the vortex chemical potentials. Moreover, because of (26) *the expansion starts with a constant term* below the KT temperature. We can now formally solve for the grand partition function in terms of the applied field

$$\mathcal{Z}[|E|] = (\sqrt{E_j E_j})^{R_0} \exp(A) \hat{\mathcal{Z}} \tag{29}$$

where $\hat{\mathscr{Z}}$ is a function of β and the chemical potentials only, while

$$A = \frac{1}{2} \frac{1}{2!} \frac{\partial^2 R}{\partial E_k \, \partial E_k}\bigg|_0 E_j E_j + \text{HOT}. \tag{30}$$

Note that $A \to 0$ when $E_j \to 0$. We are now in the position to look at the behaviour of the correlation functions.

The average momentum is simply

$$\beta \langle P_j \rangle = R_0 \frac{E_j}{E_k E_k} + \frac{\partial A}{\partial E_j}. \tag{31}$$

which *diverges* as $1/|E|$ when $E_j \to 0$ below T_{KT}. Notice, however, that $\beta \langle \delta \mathscr{H} \rangle_0 = -R_0$ is still *finite*. The fluctuations about this infinite average current in the $E_j \to 0$ limit is

$$\beta^2 \langle \tilde{P}_j \tilde{P}_k \rangle_0 = \lim_{E_n \to 0} \frac{R_0}{E_l E_l} \left(\delta_{jk} - 2 \frac{E_j E_k}{E_l E_l} \right) + \frac{1}{2} \frac{\partial^2 R}{\partial E_l \, \partial E_l}\bigg|_0 \delta_{jk}. \tag{32}$$

Taking the trace of (32) we find that *because we are working in two dimensions*

$$\rho_n = \frac{1}{2m\beta\Omega} \frac{\partial^2 R}{\partial E_l \, \partial E_l}\bigg|_0. \tag{33}$$

We see that ρ_n is finite as long as R is expandable in a Taylor series about $E_j = 0$, justifying our ansatz that $\langle \delta \mathscr{H} \rangle_0$ exists. Then, using (13), we obtain the following expression for the superfluid momentum–momentum correlation function

$$\beta^2 \varepsilon_{jl} \varepsilon_{km} \int_C \langle \tilde{g}_l^{\text{sf}}(r) \tilde{g}_m^{\text{sf}}(0) \rangle_0 \, \mathrm{d}^2 r = \lim_{E_n \to 0} \frac{n_s}{E_l E_l} \left(\delta_{jk} - 2 \frac{E_j E_k}{E_l E_l} \right) + \beta m \rho_n \delta_{jk} \tag{34}$$

where $n_s \equiv R_0 / \Omega$.

When $n_s \neq 0$ we find that the off-diagonal terms in the superfluid momentum–momentum correlation functions are infinite when $E_j \to 0$, signifying off-diagonal long range order in the static limit. Since the high-temperature phase is a normal fluid, n_s must vanish above the transition temperature T_c although it is greater than zero below it. n_s functions as an order parameter for the system. It is straightforward to show that this order parameter must also be independent of β below T_c. Calculating the total average energy of the system and using (29), we obtain the following consistency equation

$$\langle \mathscr{H} \rangle - \frac{R}{\beta} = -\frac{1}{2} \frac{\partial R_0}{\partial \beta} \log(E_j E_j) - \frac{\partial A}{\partial \beta} - \frac{1}{\hat{\mathscr{Z}}} \frac{\partial \hat{\mathscr{Z}}}{\partial \beta}. \tag{35}$$

Now let $E_j \to 0$. Whether or not $\langle \mathscr{H} \rangle_0$ is finite in this limit is determined not by how the limit is taken, but rather by the expression for $\mathscr{H}$ itself. $\hat{\mathscr{Z}}$ is independent of E_j and, from (30), $A \to 0$ as $E_j \to 0$. Because R is *finite* in the $E_j \to 0$ limit, we find that R_0 must be independent of β everywhere except, perhaps, at the transition temperature where the grand partition function itself is no longer analytic. The order parameter is therefore a step function in T with the discontinuity occurring at T_c. It is, however, still a function of the chemical potentials of the vortices.

From property 2 we conclude that $T_c \geq T_{\text{KT}}$. From our inequality (25), we note that although it may be well behaved everywhere else, there is singularity in $\langle N \rangle$ when $T = T_{\text{KT}}$. If $T_c > T_{\text{KT}}$, the system will have a singular behaviour *before* the transition temperature, which would be unphysical. On physical grounds, then, we conclude that $T_c = T_{\text{KT}}$.

There is not much more that we can say about the general ensemble. We now turn our attention to the KT ensemble. Of course all the results we have obtained so far also holds for the KT ensemble with the addition of a few new features. Using the form of the grand partition function given in (22) to calculate $\beta^2 \langle \tilde{P}_j \tilde{P}_k \rangle_0$ directly, we find that

$$R^{KT} = -2 \left(1 - \frac{\beta}{\beta_{KT}} \right) \langle N \rangle^{KT}. \tag{36}$$

The inequality is now exact since $D = N$ for the KT ensemble. Moreover,

$$4 \left(1 - \frac{\beta}{\beta_{KT}} \right)^2 \langle \tilde{N}^2 \rangle^{KT} = E_k E_k \frac{\partial^2 R^{KT}}{\partial E_l \, \partial E_l} \bigg|_0 \tag{37}$$

where $\tilde{N} \equiv N - \langle N \rangle^{KT}$ is the fluctuation in the average number of vortices. It vanishes in the static limit for all values of T except, perhaps, at the transition temperature T_{KT}. Because there are now only vortices with charge $|q| = 1$, there is only the one fugacity λ and, by using (37), we conclude that

$$\lambda \frac{\partial \langle N \rangle^{KT}_0}{\partial \lambda} = \langle \tilde{N}^2 \rangle^{KT}_0 = 0. \tag{38}$$

$\langle N \rangle^{KT}_0$ is independent of the chemical potential and from (36) we find that for the KT ensemble the order parameter n_s^{KT} is a *constant independent of all thermodynamic variables*.

We now let E_j vary slowly with time and look at how the average number of vortices in the system changes. From (33) and (37) we find that for small E_j,

$$\frac{d \langle N \rangle^{KT}}{dt} = \frac{1}{T_{KT} - T} \frac{\rho_n^{KT} \Omega}{k_B} \frac{d\mathscr{E}}{dt} \tag{39}$$

where

$$\mathscr{E} \equiv \tfrac{1}{2} m E_j E_j \tag{40}$$

is the amount of energy the external source is depositing into the system. Notice that $\langle \dot{N} \rangle^{KT}$ is coupled only to the normal fluid density and *not* to the superfluid density. Above the transition temperature the external source will decrease the average number of vortices while below the transition temperature it will tend to increase it.

Returning to (36), we find that the average density of vortices for the KT ensemble $\rho_v^{KT} \equiv \langle N \rangle^{KT}_0 / \Omega$ in the static limit is

$$\rho_v^{KT} = \frac{n_s^{KT}}{2} \frac{T}{T_{KT} - T} \theta(T_{KT} - T) \tag{41}$$

where $\theta(x)$ is the step function. Consequently, *there are no vortices in the high-temperature phase*, a result which runs contrary to the standard model of the transition and with which we shall attempt to reconcile at the end of this paper. Rather, the number of vortices is unbounded at T_{KT} and slowly *decreases* as T is lowered below T_{KT}. Because of the attractive force between oppositely charged vortices, this decrease is due to the annihilation of a charge $q = 1$ vortex with a $q = -1$ vortex. This does *not* mean, however, that the fluid is quiescent at high temperatures. In our previous work [10] we have demonstrated that the excitation spectrum of the ^{4}He film consists of two components: one a quasi-particle component corresponding to the phonon gas,

and the other a pseudo-particle, vortex gas component. In the constant density limit, the vortex gas component completely decouples from the phonon gas so that the interaction between the quantized point vortices and the quasi-particle spectrum may be neglected to lowest order. This is the regime that we are working in. Consequently, there will in fact be quasi-particle excitations in the fluid in the high-temperature phase, and these excitations may even consist of rotational flows. Our result states that these flows may *not* form ideal, point vortices with integer-valued circulation.

Let us return for a moment to the general KT ensemble. Because the energy required to create higher charged vortices is much greater than the energy needed to create vortices with unit charge, $\rho_v \lesssim \rho_v^{KT}$ where ρ_v is the average density of vortices in the general ensemble. This observation, combined with the inequality (25), gives an upper bound on the vortex density in the general ensemble

$$\rho_v \lesssim \frac{n_s}{2} \frac{T}{T_{KT} - T} \theta(T_{KT} - T) \tag{42}$$

as well as $n_s \lesssim n_s^{KT}$ (see figure 1). The average number of vortices in the general ensemble may decrease faster with temperature than the number of vortices in the KT ensemble. How much faster is still an open question, although we can obtain a formal expression for it. From (24), we find

$$\rho_v = -\frac{n_s}{2} \frac{T}{T - T_{KT}} \theta(T_{KT} - T) + \mathcal{R} \tag{43}$$

where

$$\mathcal{R} = \frac{1}{\Omega} \frac{T_{KT}}{T - T_{KT}} (\langle D \rangle_0 - \langle N \rangle_0) \tag{44}$$

is the approximate difference in the rates at which the average number of vortices decrease. Note also that because $\mathcal{R} \to 0$ as $T \to 0$, $\langle D \rangle_0 - \langle N \rangle_0 \sim (T/T_{KT})^p$ where $p > 0$.
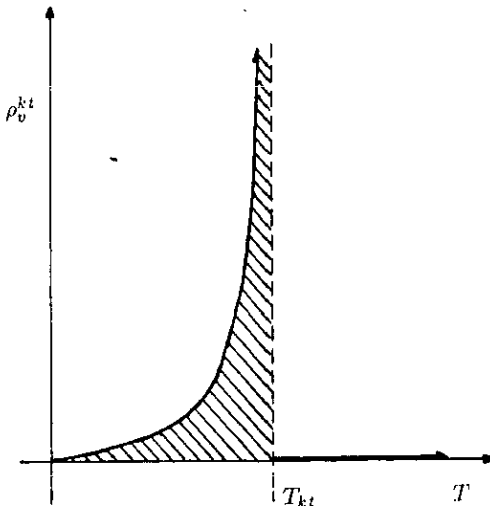
Figure 1. Graph of the density of vortices ρ_v^{KT} against T for the KT ensemble. The shaded region in the graph represents the allowed values that the density of vortices in the general ensemble ρ_v may have. Notice that there are no vortices whatsoever above the transition temperature and the singularity at $T = T_{KT}$.

5. The behaviour at $T = T_{KT}$

We now turn our attention to the behaviour of the system at the transition temperature itself. While before we fixed $\beta \neq \beta_{KT}$ and considered the behaviour of the correlation functions as $E_j \to 0$, we now fix E_j at some small but non-vanishing value and let $\beta \to \beta_{KT}^+$ ($T \to T_{KT}^-$). Our results in this section holds for both the general ensemble as well as the KT ensemble. At the transition temperature the system must be scale invariant. Kosterlitz and Thouless have shown that the relevant scale for the ensemble is the cutoff length a and to emphasize this fact we shall, in this section, denote the grand partition function as $\mathscr{Z}[a, |E|]$. Let us now scale $a \to \xi a$. Then

$$Z_N[\xi a, |E|] \prod_\alpha^N \int_C \frac{\mathrm{i}\,\mathrm{d}z^\alpha \wedge \mathrm{d}\bar{z}^\alpha}{(\xi a)^2} \exp\left(\beta e_{KT} \sum_{\alpha \neq \beta}^N q_\alpha q_\beta U\left(\left| \frac{z^\alpha - z^\beta}{\xi a} \right| \right) + \beta E_j P_j \right). \tag{45}$$

Letting $w^\alpha = z^\alpha / \xi$, we find that $\mathscr{Z}[\xi a, |E|] = \mathscr{Z}[a, \xi|E|]$ so that scaling a is equivalent to scaling the external flow field. Thus, the system must also be invariant under the scaling of E_j at the transition temperature.

Returning to (29) and using the definition of the normal fluid density, we write

$$\log \mathscr{Z}[a, |E|] = \log \hat{\mathscr{Z}} + \frac{n_s \Omega}{2} \log(E_k E_k) + \tfrac{1}{2} m\beta\rho_n \Omega E_k E_k + \text{HOT}. \tag{46}$$

Leaving the logarithmic term alone for now, we see that the normal fluid density is the coefficient of a term which has naive scaling dimension 2. As such, we conclude that $\beta\rho_n \to 0$ as $T \to T_{KT}^-$. Following Minnhagen and Warren (see note added), we define the superfluid density as $\rho_s = \rho_0 - \rho_n$, and find that

$$\lim_{T \to T_{KT}^-} \frac{\rho_s}{T} = \frac{\rho_0}{T_{KT}} = \frac{2mk_B}{\pi\hbar^2} \tag{47}$$

where k_B is Boltzmann's constant and we have used $k_B T_{KT} = e_{KT}$. This is the result originally obtained by Nelson and Kosterlitz [3] (our ρ_s is the *number* density and *not* the mass density used by Nelson and Kosterlitz, however). Moreover, we find that all the higher-order terms in the expansion of R, which are proportional to the higher order superfluid momentum correlation functions, must also vanish as $T \to T_{KT}^-$.

n_s is the coefficient of the logarithmic term in (46) which has an anomalous scaling dimension. It is unclear exactly what, if any, restrictions scale invariance will impose upon it. If, however, we take the scaling invariance to also include the logarithm, we find that $n_s(T \equiv T_{KT}) = 0$ which substantiates our earlier physical arguments for the vanishing of n_s above the transition temperature in section 4.

6. Conclusion

From the above results we are lead to a somewhat different interpretation of the KT superfluid phase transition. Approaching the phase transition from above we see that *there are no vortices whatsoever* above the transition temperature. From the HMW theorem the ^{4}He film has no choice but to behave as a normal fluid. As T is lowered below the KT temperature, however, the vacuum state becomes unstable and any external perturbation of the system causes a gas of vortices to be *created*. Due to

angular momentum conservation, this gas must be neutral and consequently the ^{4}He film behaves as a superfluid. The HMW theorem is circumvented due to the *presence* of vortices in the fluid. As the temperature is lowered further, the average number of vortices in the system decreases due to pairwise annihilation of oppositely charged vortices until there are no vortices left at $T = 0$. Approaching the transition temperature from below, the increasing temperature creates vortices *pairwise* in the fluid to preserve neutrality. The average separation between vortices now increases with temperature, thereby decreasing the likelihood of pair annihilation. This separation, however, is bounded above by the *finite* size of the system while the average number of vortices will continue to grow without bound. At the transition temperature itself a massive annihilation of the vortices will occur so that no vortices are left above T_{KT}.

At first glance this seems to be different from the standard model of the phase transition which has free vortices above the transition temperature and vortices bound in dipole pairs below it. We note, however, that in the KT theory two vortices are considered to be bounded when they come within a distance a, the cut-off length, of each other. From our point of view they have effectively annihilated each other. What Kosterlitz and Thouless interpreted as vortex binding and unbinding may also be viewed as vortex annihilation and creation. Our ρ_v^{KT} is, in the language of the standard model, the density of free vortices in the liquid and long range order is destroyed by these vortices when the number of vortices in the system becomes infinite in a finite size system. Aside from the requirement that there are no vortices in the high-temperature phase, there is little difference between our theory and the KT theory.

Although we have allowed E_j to vary with time, ours is not a complete theory of the *dynamical* phase transition because the vortices have not been coupled to the underlying fluid. E_j is treated as an *external* field and has not been related to any of the elementary excitations in the ^{4}He film. The reader is referred to either [13] or [14] for the complete description of the dynamical phase transition.

As we have seen, the magnitude of the external field $|E|$ plays an analogous role to the chemical potential for the system. It enters, however, into the grand partition much in the same way as the external magnetic field does for the two-dimensional Ising model: as a Lagrange parameter. Identifying $|E|$ as the chemical potential, equation (29) shows that when $n_s > 0$ there is a zero of the partition function at $|E| = 0$. This is reminiscent of the Yang-Lee theory of phase transitions in which a phase transition occurs when the partition function develops a zero as the complexified fugacity z pinches the real axis at $z = 1$ [15, 16]. The crucial difference, of course, is that E_j is a vector, and not a scalar like the chemical potential.

Note added. There are two different definitions of the superfluid density currently in use in the literature. The one used by Nelson and Kosterlitz is

$$\left(\frac{\rho_s}{\rho}\right)_{NK} = \left(1 + \frac{\beta}{m}\int_\Omega \langle g^{sf}(r) \cdot g^{sf}(0)\, \mathrm{d}^2 r\rangle\right)^{-1} \tag{N1}$$

while the one given by Minnhagen and Warren is

$$\left(\frac{\rho_s}{\rho}\right)_{MW} = 1 - \frac{\beta}{m}\int_\Omega \langle g^{sf}(r) \cdot g^{sf}(0)\, \mathrm{d}^2 r\rangle \tag{N2}$$

where we have used our definition of g^{sf} against Minnhagen and Warren's (they differ by a factor of $\rho h/m$). Minnhagen and Warren have shown that the two definitions are equivalent near T_{KT}. Our definition of ρ_s agrees with Minnhagen and Warren's once the difference of the factor of two in the definitions of the vortex Hamiltonian (8) is taken into account. All three definitions of the superfluid density give the same result at T_{KT}, namely (47).

Appendix

Due to the importance of the relationship between the polarization vector and the superfluid velocity, we shall establish the identity (12). We begin with a very brief discussion of differential forms. The reader is referred to [17] for a complete description.

The object $\wedge$ in the integration measure $i\,dz \wedge d\bar{z}$ defines a *wedge product*. Its usefulness will be made clear later on. Let f be any function of the coordinates x_j. The *exterior derivative* df of f is defined as

$$df \equiv \frac{\partial f}{\partial x_j}\,dx_j \tag{A1}$$

with the rule $dx_j \wedge dx_k = -dx_k \wedge dx_j$. In evaluating the integral in (12), we find it much more convenient to use polar coordinates and we let $z = r\exp(i\theta)$. Then from (A1), $dz = (dr + ir\,d\theta)\exp(i\theta)$ so that $dz \wedge d\bar{z} = -2ir\,dr \wedge d\theta$ which is obtained without the use of the Jacobian of the transformation. Then

$$\frac{1}{2\pi i}\int_{\mathscr{D}} \frac{dz \wedge d\bar{z}}{z - z^\alpha}$$

$$= \frac{1}{\pi z^\alpha}\int_0^{|z^\alpha|}\int_0^{2\pi}\left(1 - \frac{r\,e^{i\theta}}{z^\alpha}\right)^{-1} r\,dr\,d\theta - \frac{1}{\pi}\int_0^{R}\int_0^{2\pi} e^{-i\theta}\left(1 - \frac{z^\alpha\,e^{-i\theta}}{r}\right)^{-1} dr\,d\theta \tag{A2}$$

where R is the radius of the disk $\mathscr{D}$. We have broken the integral up into two pieces so as to make use of the expansion

$$\frac{1}{1 - \bar{z}} = \sum_{n=0}^{\infty} z^n \tag{A3}$$

which holds as long as $|z| < 1$. Because

$$\frac{1}{2\pi}\int_0^{2\pi} e^{in\theta}\,d\theta = \delta_{n,0} \tag{A4}$$

we find that

$$\frac{1}{2\pi i}\int_{\mathscr{D}} \frac{dz \wedge d\bar{z}}{z - z^\alpha} = \frac{2}{z^\alpha}\int_0^{|z^\alpha|} r\,dr = \bar{z}^\alpha \tag{A5}$$

and we are done.

References

[1] Kosterlitz J M and Thouless J D 1973 *J. Phys. C: Solid State Phys.* **6** 1181
[2] Kosterlitz J M 1974 *J. Phys. C: Solid State Phys.* **7** 1046
[3] Nelson D R and Kosterlitz J M 1977 *Phys. Rev. Lett.* **39** 1201
[4] Bishop D J and Reppy J D 1978 *Phys. Rev. Lett.* **40** 1727
[5] Berezinskii V L 1971 *Zh. Eksp. Teor. Fiz.* **61** (1972 *Sov. Phys.–JETP* **34** 610)
[6] Hohenberg P C 1967 *Phys. Rev.* **158** 383
[7] Mermin N D and Wagner H 1966 *Phys. Rev. Lett.* **17** 1133
[8] Minnhagen P and Warren G G 1981 *Phys. Rev.* B **24** 2526
[9] Friedrichs K O 1966 *Special Topics in Fluid Dynamics* (New York: Gordon and Breach) ch 19
[10] Speliotopoulos A D and Morrison H L 1989 *Phys. Lett.* **141A** 284
Morrison H L, Albertin U K and Lindesay J V 1985 *Phys. Lett.* **108A** 8

[11] Chapman D M F 1978 *J. Math. Phys.* **19** 1988
[12] Speliotopoulos A D and Morrison H L 1991 *J. Math. Phys.* in press
[13] Ambegaokar V, Halperin B I, Nelson D R and Siggia E D 1978 *Phys. Rev. Lett.* **40** 783
[14] Ambegaokar V, Halperin B I, Nelson D R and Siggia E D 1980 *Phys. Rev.* B **21** 1806
[15] Yang C N and Lee T D 1952 *Phys. Rev.* **87** 404
[16] Lee T D and Yang C N 1952 *Phys. Rev.* **87** 410
[17] Flanders H 1963 *Differential Forms, with Application to the Physical Sciences* (New York: Academic)