COURSE 9

INTRODUCTION TO TOPOLOGICAL QUANTUM NUMBERS

D.J. THOULESS

Dept. of Physics, Box 351560,
University of Washington, Seattle,
WA 98195, U.S.A.

Contents

INTRODUCTION TO TOPOLOGICAL QUANTUM NUMBERS

D.J. Thouless

Preface

These lecture notes were prepared rather soon after I completed my book on *Topological quantum numbers in nonrelativistic physics*, which was published by World Scientific Publishing Co. Pte. Ltd., Singapore, in early 1998. I have not attempted to make a completely fresh presentation, but have cannibalized the text of my book to produce something shorter, with a different ordering of topics. I wish to thank the publishers for allowing me to do this self-plagiarization.

1 Winding numbers and topological classification

1.1 Precision and topological invariants

High precision work generally depends on two ingredients. These are reproducibility, and the reduction of a measurement to a counting procedure. A ruler is a device for comparing a length with the number of marks along the ruler, and a vernier allows interpolation between marks on the main scale also to be done by counting. A pendulum clock and its successors are devices for comparing a time interval with the number of ticks that occur in the interval. Such devices are not completely reproducible, and may vary when conditions change. The earth's rotational and orbital motion provide time standards that can be used for calibration, but they are difficult to measure with very high precision, and we know that the rotational motion is subject to random as well as to systematic changes. Cesium atoms and ammonia molecules are reproducible, and they can form the basis for length measurements in which interference fringes are counted, or as time standards by driving the system in resonanance with with a standard atomic or molecular transition and counting beats against some uncalibrated frequency.

The measurement of g for the electron to parts in 10^{11} is achieved, in part, by measuring the frequency difference between the spin and twice the orbital resonance frequency in a magnetic field.

Counting can be made very precise because, although any particular counter may make mistakes, comparison between the outputs of several independent counters can reduce the error rate to an extremely low value.

Over the past 30 years several devices have been developed for use in high precision work where the devices themselves are manifestly not reproducible, but nevertheless give fantastically reproducible results. Among such devices are the SQUID magnetometer, which compares magnetic flux in a superconducting ring with the quantum of flux $h/2e$ for a superconductor, the Josephson voltmeter, which compares the frequency of a microwave device with the frequency $2\ eV/h$ generated by a voltage V across a superconducting weak link, and a quantum Hall conductance standard, which compares electrical conductance with the quanta e^2/h for conductance in a quantum Hall device. In none of these cases does the fabrication of the device have to be very tightly controlled, but there are good theoretical reasons and very strong experiments to show that different devices, even those made of different materials, give measurements that are essentially identical. This insensitivity to details is a characteristic of topological quantum numbers that is one of the themes of my next three lectures.

We are used to thinking of quantum numbers like angular momentum which are related to invariance principles, and which can be studied from the algebra of the generators of the symmetry group. Such quantum numbers are sensitive to breaking of the symmetry, and are generally not useful in environments that are poorly controlled, such as interfaces between solids.

1.2   Winding numbers and line defects

The simplest type of topological quantum number that I discuss is the winding number of an angle such as the phase of a condensate wave function in a superfluid or superconductor. If a neutral superfluid with a complex scalar order parameter, such as superfluid ^{4}He, is contained in the interior of a torus, or if a superconductor is made in the form of a ring, the condensate wave function has the form

$$\Psi(\mathbf{r}) = |\Psi| \exp(iS) . \tag{1.1}$$

Single-valuedness of the condensate wave function Ψ implies that the phase S is locally single valued, but it may change by a multiple of 2π on a closed path that goes round the hole in the middle of the ring. The winding number

$$W = \frac{1}{2\pi} \oint \mathbf{grad} S \cdot \mathbf{dr}, \tag{1.2}$$

where the path of the integral is a simple loop round the hole, is an integer. This just depends on the topology of the region containing the superfluid or superconductor, and on the nature of the order parameter, and is independent of detailed geometrical features such as symmetry, or of detailed material properties such as homogeneity.

I have given a description of the mathematical definition of a winding number, but have not yet said why it is relevant to physics. In the case of a neutral superfluid equation (1.2) defines the quantum number associated with superfluid circulation, whose quantization was argued by Onsager [1]. This is a quantity that can be measured, and this was first done by Vinen [2]. Measurement of the circulation of a neutral superfluid is difficult, and there are some real problems with its definition. The winding number for a superconductor counts the number of quanta of magnetic flux, and this can be measured with very high precision, thanks to the Josephson effects.

Winding numbers are not only of importance in nonsimply connected systems, such as the interior of a torus. In superfluids and superconductors the order parameter can go to zero along curves that either run across the system or form closed loops within the system. Around such line singularities there may be nonzero integer winding numbers. Quantized vortex loops were detected in superfluid ^{4}He by Rayfield and Reif [3]. In Type II superconductors flux lines are line singularities which carry one quantum of magnetic flux.

A nonzero winding number assures the topological stability of the line singularity. Consider some small cylindrical region with its axis close to part of a curve along which the order parameter goes to zero. If the winding number is nonzero around some loop on the boundary of the region, there is no continuous change of the order parameter which can be made that will remove the singularity, since the winding number, with integer values, cannot be changed continuously, and must remain nonzero. If the winding number is zero the interior of the cylinder can be filled in with a continuous order parameter which is nonzero everywhere inside, so that the line singularity has two disconnected ends. This cylinder can then be expanded in the perpendicular along the line defect in such a way that the whole defect is replaced by a continuous nonzero value of the order parameter.

In 1931, Dirac [4, 5] gave an argument for the quantization of electric charge in which a similar winding number appears, essentially the integral of the magnetic vector potential round a closed circuit. This winding number appears in the phase of the wave function when you try to write a quantum theory for an electrically charged particle in the presence of a magnetic monopole. The singularitiy that the circuit encloses is not a physical singularity, but a line singularity of the vector potential in the chosen gauge. Although magnetic monopoles have not been experimentally detected, and may not exist, there is no doubt that electric charge is quantized with a

very high precision. Already in 1925 Piccard and Kessler [6] had shown that the charge of CO_2 molecules was sufficiently low that electrons and protons must have charges whose magnitudes differ by less than one part in 10^{20}, and later work has improved on this bound by more than an order of magnitude [7].

The quantum Hall effect provides another example of a topological quantum number which corresponds to a physical variable that can be measured with high precision – the Hall conductance. In this case the simplest system on which a measurement can be made is a two-dimensional electron gas with two pairs of leads – one pair to pass the current through the system, and another pair to measure the voltage. Such a circuit has the topology of a torus with a single hole in it, as I discuss later. It turns out that the Hall conductance can be related to a Chern number, a number associated with a torus, rather than a winding number associated with a loop. This is another example of a topological quantum number which is quantized with high precision, although the precision has not been established as accurately as flux quantization and charge quantization have been.

1.3   Homotopy groups and defect classification

For systems in which the line defects are characterized by winding numbers, such as superfluids characterized by a complex scalar order parameter, or magnets with a preferred plane of magnetization, there is an obvious way of combining defects. The winding number round a path that encloses several defects is the algebraic sum of the winding numbers of associated with each of the defects that goes through a simply connected surface that is bounded by this path. One can easily show this by continuously deforming the path until it is broken up into a sum of loops which each contain only one singularity. Conversely one can combine paths round individual defects, expanding them in a continuous way, and gluing together the separate paths to form a big path surrounding more than one defect. The algebra of this process is just the algebra of addition of signed integers.

The assignment of a phase angle to each point on a path round one or more singularities can be regarded as a mapping of the loop onto a circle. Such a mapping has a single topological invariant, the winding number. The law of combination of mappings of different loops, by the process of continuous deformation and gluing together, defines a group, the first homotopy group π_1, which in this case is just the group Z of integers.

Other types of order parameter can give different homotopy groups. For the Heisenberg ferromagnet, in which the magnetization can be oriented in any direction, the mapping of a loop surrounding a line defect will be onto the surface of the 2-sphere corresponding to the possible directions of magnetization. The mapping of a loop onto the surface of a sphere can

always be contracted to a single point, so the homotopy group π_1 in this case is just the trivial group which has only the identity element.

In the A phase of superfluid ^{3}He and in some of the liquid crystal phases the homotopy groups are finite, so there are only a finite number of different topological states of such systems. For example, for a uniaxial nematic liquid crystal, with the order parameter specified by a director on the surface of a projective sphere (a sphere in which opposite points are equivalent) rather than by a vector on the surface of a sphere, there is a topological invariant for a system confined to the interior of a torus. This invariant takes the value zero if the director passes through the equator of the projective sphere an even number of times on a path round the system, and unity if the director passes through the equator of the projective sphere an odd number of times. The homotopy group π_1 in this case is just the group Z_2 with two members. For biaxial nematics and for cholesterics the symmetry group of the order parameter is noncommutative, and this gives some extra complications to the theory of defects in such materials.

For these systems in which the homotopy group is finite we cannot expect a measurable physical variable to be quantized in the way that circulation in ^{4}He and flux in superconductors are quantized. The superfluid phases of liquid ^{3}He illustrate this point. Whereas circulation in the B phase is quantized in much the same way as the circulation in ^{4}He, there is no quantized circulation in the A phase. There is a topological quantum number, but it has only the values 0 and 1, and does not correspond in any direct way to the circulation of the fluid. Such quantum numbers are useful for classifying defects, and for determining whether two apparently different states of the system can actually be continuously changed from one to the other.

I use the term topological quantum numbers regardless of whether the topological invariant actually has anything to do with quantum mechanics. I have already mentioned the examples of magnets and liquid crystals, where the order parameters have little connection with Planck's constant. A crystalline solid is another case of a system with topological quantum numbers. There are two important order parameters in a solid, which are the position of the actual unit cell with respect to an ideal unit cell, and the orientation of the unit cell. Long range correlation of the positional order is required for the observation of the sharp Bragg peaks which are measured in an X-ray diffraction experiment – the Debye-Waller factor gives the reduction in magnitude of this long range order. Elasticity theory deals with the effects of slow modulation of positional and orientational order. Crystal dislocations are the topological defects associated with the positional order. Orientational changes are less important in solids than positional changes, because they are costly in the elastic energy associated with the accompanying changes in positional order, but disclinations are the defects associated with orientational order.

For defect surfaces and for point defects there are other types of homotopy groups π_0 and π_2 that determine the topological stability of the defects. If the order parameter is singular everywhere on a surface, then it is singular at some point on any short line segment that crosses the surface. Such a singularity on a line segment can only be topologically stable if the order parameter has two disconnected values on the two sides of the surface. The homotopy group π_0 is therefore nontrivial only when the order parameter has values lying in distinct regions, as it does in the Ising model of magnetism, or in the Potts model. In some of Onsager's early discussions of the circulation of a superfluid [8] the idea was suggested that regions of different quantized circulation should be nested inside one another, separated by singular surfaces — vortex sheets. Such vortex sheets are topologically unstable, as the order parameter can be made continuous across the sheet in small regions, which then can be expanded until the vortex sheet has broken up into an array of vortex lines. Only when the order parameter space breaks up into distinct regions, which it does in the A phase of ^{3}He [10,11], can a vortex sheet be stable.

A point defect can be surrounded by a spherical surface, and the behavior of the order parameter on this sphere defines the homotopy group π_2 which describes the topological properties of a point defect. If the order parameter is an angle the order parameter space is a circle. All continuous maps of a 2-sphere onto a circle are trivial, and can be shrunk to a single point, so π_2 for superfluid ^{4}He or for a planar magnet is trivial. For the Heisenberg model the direction of the order parameter lies on a sphere, and the mapping of one sphere onto another can be characterized by the topological invariant known to Euler

$$
\begin{aligned}
N_{\text{w}} &= \frac{1}{4\pi} \int_0^{2\pi} \mathrm{d}\phi_{\text{s}} \int_0^{\pi} \mathrm{d}\theta_{\text{s}} \sin\theta_{\text{d}} \frac{\partial(\theta_{\text{d}}, \phi_{\text{d}})}{\partial(\theta_{\text{s}}, \phi_{\text{s}})} \\
&= \frac{1}{4\pi} \int_0^{2\pi} \mathrm{d}\phi_{\text{s}} \int_0^{\pi} \mathrm{d}\theta_{\text{s}} \hat{\mathbf{d}} \cdot \frac{\partial \hat{\mathbf{d}}}{\partial \theta_{\text{s}}} \times \frac{\partial \hat{\mathbf{d}}}{\partial \phi_{\text{s}}} ,
\end{aligned}
\tag{1.3}
$$

where $\hat{\mathbf{d}}$ is the direction of the order parameter and $\theta_{\text{d}}, \phi_{\text{d}}$ its polar angles, while $\theta_{\text{s}}, \phi_{\text{s}}$ are the polar angles of points on the sphere relative to the point defect. This quantity has integer values, and an example of a defect whose quantum number is $+1$ is shown in Figure 1. Point defects of this sort can be combined according to the rules of integer addition. A defect with quantum number -1 would be obtained by reversing the direction of the magnetization everywhere.

In addition to these topologically stable line and point defects there is considerable interest in extended defects, domain walls, textures or solitons, which can also be characterized by a topological quantum number. A domain wall between two oppositely aligned domains of a ferromagnet is
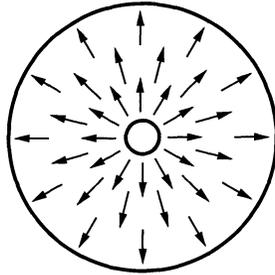
Fig. 1. Magnetization pointing outwards in the space between two spherical enclosing surfaces. This is known as a hedgehog.

an example of such a structure, and a number of other cases occur in the context of superfluid ^{3}He and liquid crystals. Such textures may be stabilized by the effect of boundary conditions, although they are not topologically stable defects. For example, if the magnetization at two ends of an ideal isotropic magnetic bar are constrained to point in different directions there will be a domain wall dividing the two different directions of magnetization.

2   Superfluids and superconductors

2.1   Quantized vortices and flux lines

The superfluid component of a neutral superfluid is supposed to flow with a velocity determined by the phase S of the condensate wave function, so that for ^{4}He the superfluid velocity is

$$\mathbf{v}_\mathrm{s} = \frac{\hbar}{m_4}\mathbf{grad}S\,, \tag{2.1}$$

where, if the flow is incompressible, S satisfies the Laplace equation. The flow is therefore potential flow, except where there are singularities. The normal component can be envisaged as a gas of excitations moving in the medium determined by the condensate wave function. On a large scale the motion of superfluid helium is not really irrotational even when the density of the normal component is very small. Rotation of a beaker of superfluid brings the surface into much the same parabolic shape that rotation of a normal fluid would produce. This led Onsager [1] to argue that the curl of the superfluid velocity should be concentrated into singular lines – quantized vortices.

Since the phase S of the order parameter has to be single-valued modulo 2π, the circulation round one or more of these singularities must be a

multiple of $\kappa_0 = h/m_4$:

$$\oint \mathbf{v}_\mathrm{s} \cdot \mathrm{d}\mathbf{r} = \frac{\hbar}{m_4} \oint \mathbf{grad}\, S \cdot \mathrm{d}\mathbf{r} = \frac{2\pi n\hbar}{m_4} = n\kappa_0 \,, \qquad (2.2)$$

which is just the form given in equation (1.2). Rigid body rotation with angular velocity ω is simulated by a density $2\omega/\kappa_0$ per unit area of these quantized vortices. At a speed of one revolution per second this leads to a concentration of vortex lines of about 1.25 mm^{-2}.

These vortex lines behave in many ways like the vortices of classical hydrodynamic theory, and much of the relevant theory can be found in Lamb's book on *Hydrodynamics*. Vortices are carried along by the local potential flow velocity, produced generally by a combination of externally imposed flow and the flow produced by vortex lines themselves. Vortex lines have an energy per unit length which is $\rho\kappa^2$ times a logarithmic factor, where ρ is the mass density of the fluid and κ is the circulation round the vortex. The argument of the logarithm is the ratio of two length scales, a large length which is the size of the container or the distance between vortices, and a small distance cut-off a_0 which gives the size of the vortex core, the distance at which potential flow no longer occurs. Straight vortices in an infinite container, or along the axis of a cylindrical container, are stable, and have normal modes of circularly polarized vibration. Circular vortex rings of radius R are also stable, and propagate in a direction perpendicular to their planes at a speed proportional to κ/R, also multiplied by a logarithm of R/a_0.

In Feynman's description the wave function Ψ_V for a system with a vortex line centered on the axis of cylindrical coordinates r, ϕ, z has the form

$$\Psi_\mathrm{V} \approx \exp\left[in \sum_j \phi_j - \sum_j \alpha(r_j) \right] \Psi_0 \,, \qquad (2.3)$$

where Ψ_0 is the ground state. The velocity is $n\hbar/m_4 r$ in the azimuthal direction. The factor $\exp[-\alpha(r)]$ is designed to reduce the density near the vortex core where the velocity is high. Similar results are obtained in the work of Pitaevskii [12] and of Gross [13], who studied a nonlinear Schrödinger equation in detail. In most models the energy is lowest if the vortices carry a single quantum of circulation, so n is ± 1. In general the vortex will not be straight, but will follow some curving, time-dependent path, but in principle the theories can be modified to allow for this, although the practical difficulties are great.

The phase S of the order parameter will change by a multiple of 2π when it goes round a closed path that surrounds vortex lines. The order

parameter goes to zero on the vortex lines themselves, where the topological singularity resides. As I have argued in Section 1, such a singularity is topologically stable. This quantization of circulation is sometimes regarded as a manifestation of a quantization of angular momentum of $\hbar$ per helium atom, but this is not appropriate, as the quantization of circulation is much more robust than quantization of angular momentum.

Quantized vortex lines not only solve the paradox of apparent rigid body rotation in terms of a uniform density of vortex lines, but they play many other important roles. Phonons and rotons do not transfer energy directly to or from the superfluid, but they can scatter off the vortex lines, giving up energy and momentum to vortex waves, so that there is mutual friction between the normal fluid and the vortices. This was used by Hall and Vinen [14, 15] to detect the vortex array.

Vortices provide one plausible mechanism for relaxation of superfluid flow around a ring. A direct transition between two states with different circulation would have an incredibly small matrix element, but a unit of circulation can be lost in a continuous manner by a vortex moving from one side to the other of the ring, in the manner shown in Figure 2. The energy and the average circulation are both reduced as the vortex moves in such a direction that its own circulation enhances the flow velocity ahead of it and reduces the flow velocity behind it. In general there will be a barrier to the initial formation of a vortex, which must be overcome by thermal activation or by quantum tunneling, but then the vortex moves across the circulating superfluid, losing energy to the normal fluid. This is discussed in work by Vinen [16], Langer and Fisher [17], Muirhead et al. [18], and in the book by Donnelly [19]. The other plausible method for decay of circulation is by phase slip, in which the magnitude of the order parameter temporarily goes to zero on a cross-section of the ring, and the phase slips by a multiple of 2π across that cross-section.

It was suggested by Onsager [1], and later by Feynman [20], that the λ-transition from superfluid to normal fluid might be due to the thermal nucleation of indefinitely long vortices rather than to the complete destruction of the order parameter. This is not widely believed to be correct for bulk helium, because the superfluid transition seems to have critical exponents close to those expected for a standard planar spin model in which the order parameter goes continuously to zero at the critical point. However, a phase transition analogous to the one suggested by Onsager and Feynman occurs in helium films, as is discussed in Section 5.

Finally, vortex cores are regions of low pressure and density, and they act as sites for the trapping of impurities such as ions and ^{3}He atoms. Ions trapped on vortices provide some of the best tools for studying vortices. The trapping of ^{3}He atoms has important consequences for the energetics of vortex nucleation, since such impurities usually occur in helium unless care
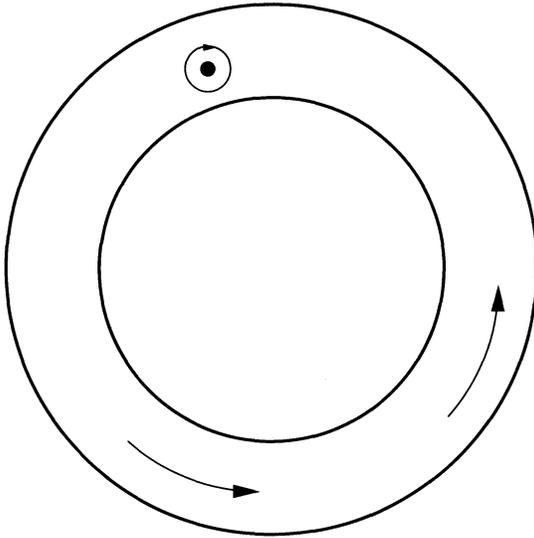
Fig. 2. A vortex moving across a persistent current. When it crosses the ring (from the outside to the inside in the case shown) it reduces the circulation round the ring by κ_0.

is taken to remove them, and the trapped impurities lower the nucleation energy for vortices.

The B phase of superfluid ^{3}He is similar to superfluid ^{4}He in many ways. Like the superconductors which are discussed below, the condensate is made up of weakly bound pairs of fermions, so the transition temperature is much lower, and there is a long correlation length. The pairs are in a triplet P state, but in the B phase spin and orbital angular momenta are combined in such a way that the square modulus of the gap parameter is isotropic. One possible way of doing this is to form a $J = 0$ combination of the orbital and spin angular momentum, but relative rotation of spin and orbital space does not affect the pairing energy, and only changes the very small hyperfine energy. If this variable orientation between spin and orbital axes is ignored the order parameter is essentially a complex scalar, and the superfluid velocity is proportional to the gradient of its phase. Circulation is therefore quantized, with a quantum of circulation $h/2m_3$, since the basic units are pairs of atoms of mass m_3.

There are many parallels between superfluidity in liquid ^{4}He and superconductivity in metals, but there are also some important differences. The order parameter is a complex scalar in both cases, so there are analogs to

quantization of circulation and to vortex lines. The differences come under three heads:

1. Electrons are fermions while ^{4}He atoms are bosons, so that the Bose condensation that occurs in superconducting metals is of very weakly bound electron pairs, rather than of the well-separated bosons of liquid ^{4}He.

2. Electron pairs are charged, so there is an important coupling with the electromagnetic field.

3. Electron pairs are in intimate contact with the periodic potential formed by the lattice of positive ions, and also with the impurities in that lattice.

For some other systems these differences can come in different combinations. For example, liquid ^{3}He and neutron stars give fermion pairing without electric charge or background effects (lattice and fixed impurities). In thin helium films the disorder of the substrate potential is important. In thin film superconductors the background potential is important, but the electric charge is much less important.

Understanding of superconductors is based on the BCS (Bardeen, Cooper and Schrieffer) theory of superconductivity [21], which involves a sort of Bose condensation of pairs of electrons whose binding energy is much less than the Fermi energy. In the standard BCS superconductor the pair is in a rather large ($\approx$100 nm) singlet state, and the pairing energy is a very small fraction of the Fermi energy, typically a millivolt or less. The core of a vortex behaves in some ways like a normal metal, although there is an energy gap of order Δ^2/E_F in the core, where Δ is the energy gap of the superconductor and E_F is the Fermi energy. This is very small for conventional superconductors.

The electric charge is very important, and leads to a number of consequences. Firstly, currents in superconductors are easy to detect, because they produce magnetic fields, whereas supercurrents are hard to detect in superfluid ^{4}He. As a result of this, very accurate measurements can be made. It was known very early to Kammerlingh-Onnes that supercurrents in a metal ring can have a negligible decay rate, whereas neutral superfluid flow escaped detection for twenty five years.

The most important result of the coupling of the electromagnetic field to the order parameter from the theoretical point of view is the Meissner effect – magnetic flux is expelled from superconductors. The Meissner effect was explained by London in terms of the rigidity of the superconducting wave function. Since the current density operator is

$$(e/m)(i\hbar\mathbf{grad} - 2e\mathbf{A}),\qquad(2.4)$$

a rigid wave function that does not respond to changes in the vector potential gives a current density equal to

$$\mathbf{j} = -\left(\frac{e\hbar}{2m}\mathbf{grad}\,S + \frac{e^2}{m}\mathbf{A}\right)n_{\mathrm{s}}, \qquad (2.5)$$

where n_{s} is the superconducting electron density, S is the phase of the superconducting wave function, and $n_{\mathrm{s}}/2$ is its square modulus. The curl of this equation gives

$$\mathbf{curl}\,\mathbf{j} = -\frac{e^2 n_{\mathrm{s}}}{m}\mathbf{B}\,. \qquad (2.6)$$

The curl of this equation, combined with Ampère's law $\mathbf{curl}\,\mathbf{B} = \mu_0\mathbf{j}$ and the charge conservation law $\mathbf{div}\,\mathbf{j} = 0$, gives the London equation

$$\nabla^2\mathbf{j} = \frac{e^2\mu_0}{m}n_{\mathrm{s}}\mathbf{j}. \qquad (2.7)$$

This shows that the current density $\mathbf{j}$ has exponential decay over a distance

$$\lambda_L = \sqrt{\frac{m}{\mu_0 e^2 n_{\mathrm{s}}}}, \qquad (2.8)$$

the London penetration depth. The conclusion of this argument is that all supercurrents are concentrated into the surface of a superconductor, and the current density in the interior of a sample falls off exponentially with the ratio of the distance from the surface to the London penetration depth. From equation (2.6) it is clear that the flux density is also zero in the interior of the superconductor, and the Meissner effect is obtained.

This gives a good description of a strongly Type II superconductor. For a Type I superconductor the same qualitative effects occur, but there is more adjustment of the condensate wave function to the magnetic field, so that the expression (2.8) for the penetration depth is altered.

Multiplication of equation (2.5) by $m/e^2 n_{\mathrm{s}}$ and integration round a closed loop inside the superconductor gives

$$-n\frac{h}{2e} = \frac{m}{e^2 n_{\mathrm{s}}}\oint \mathbf{j}\cdot\mathrm{d}\mathbf{R} + \oint \mathbf{A}\cdot\mathrm{d}\mathbf{R} = \frac{m}{e^2 n_{\mathrm{s}}}\oint \mathbf{j}\cdot\mathrm{d}\mathbf{R} + \iint \mathbf{B}\cdot\mathrm{d}\mathbf{S}\,. \quad (2.9)$$

The quantity on the right hand side of this equation is known as the fluxoid, and it is the quantized quantity. Deep in the interior of the superconductor the current density is zero, apart from terms exponentially small in the ratio of the depth in the sample to the penetration depth, and so the flux through a surface whose edge lies well inside the superconductor is $h/2e$ times the winding number of the phase. In the case of low density n_{s} of superconducting electrons, where the penetration depth is very large, the first term on the right hand side of this equation dominates, and it reduces

to the quantization of circulation for a neutral superfluid given by equation
(2.2). This result, without the factor of 2 in the electron pair charge, was
obtained by London [22].

This expression, like the expression for the circulation of a neutral super-
fluid, still involves the winding number of the phase of a condensate wave
function. However, it results from an integral round a loop of the canonical
momentum density, so that the integrand is gauge-dependent, although the
integral is not. Where the integral of the current density is negligible, which
it is inside the superconductor at depths greater than a small multiple of the
penetration depth, except for Type II superconductors in fields greater than
B_{c1}, the quantized quantity is magnetic flux, and this is something that can
be measured far more readily than the circulation of a neutral fluid.

In Type I superconductors magnetic flux is completely expelled for weak
fields, and when it begins to penetrate the positive interface energy between
the flux-free superconductor and the flux-carrying normal metal causes the
magnetic field to be aggregated in domains carrying many quanta of flux.
For a Type II superconductor the magnetic field begins to penetrate the
superconductor at fields above B_{c1}, and the negative interface energy favors
singly quantized flux lines each carrying flux $h/2e$. It is this mixed state of
the Type II superconductor, with uniformly spaced flux lines, that closely
resmbles a rotating superfluid with a uniform array of singly quantized
vortex lines.

2.2 Detection of quantized circulation and flux

There are three techniques that have been used to show quantized circula-
tion and the properties of quantized vortices in superfluid ^{4}He directly. The
first was developed by Vinen [2]. In this experiment there is a straight wire
under tension along the axis of a cylinder filled with liquid helium. The
helium is set into rotation by initially rotating the whole system above the
λ-point, cooling the helium through the transition to the superfluid state,
and then bringing the apparatus to rest, leaving the superfluid circulating
around the wire. The circulation is measured by using the Magnus force
that the circulating superfluid exerts on the wire. A derivation of this is
given in the book by Putterman [23]. The component of the force transverse
to the direction of motion has the form, very similar to the form known from
classical hydrodynamics,

$$\mathbf{F}_M = \int \rho_s \kappa_s (\mathbf{v}_L - \mathbf{v}_s) \times d\mathbf{l}_L + \int \rho_n \kappa_n (\mathbf{v}_L - \mathbf{v}_n) \times d\mathbf{l}_L , \qquad (2.10)$$

where ρ_s, ρ_n are the superfluid density and normal fluid density respec-
tively, $\mathbf{v}_s$, $\mathbf{v}_n$ are the velocities with which the superfluid and normal fluid
components are flowing past the wire, $\mathbf{v}_L$ is the velocity of the wire, $d\mathbf{l}_L$

an element of its length, and κ_s, κ_n are the circulations of the two components around the wire. The superfluid circulation should have its quantized value $\kappa_0 = h/m$, while normal fluid viscosity should reduce the normal fluid circulation to zero, so the expected form of the Magnus force would be

$$\mathbf{F}_M = \int \rho_s \kappa_0 (\mathbf{v}_L - \mathbf{v}_s) \times d\mathbf{l}_L . \tag{2.11}$$

In addition there is a dissipative force in the direction of motion if ρ_n is nonzero. There is, however, dispute about the correct form of the Magnus force when the normal fluid density is non-negligible; a recent summary of the situation has been given by Sonin [24].

The Magnus force breaks the degeneracy of the two fundamental vibrational modes of the wire, giving a splitting of the circularly polarized modes proportional to $\rho_s \kappa_0$, so that a direct measure of this quantity is made. Vinen's experiments showed a splitting that agreed with the expected value with a precision of about 3%.

In one later version of this experiment, Whitmore and Zimmermann [25] worked at relatively high temperatures, where ρ_n is significant, and confirmed that the transverse force is proportional to ρ_s. Zieve et al. [26] recently repeated Vinen's experiment, with somewhat higher precision, and used the ^{4}He measurements as a calibration for a similar experiment on the B phase of superfluid ^{3}He [27], for which they confirmed that the quantum of circulation is indeed $h/2m_3$.

Rayfield and Reif [3, 28] used the trapping of ions on vortex rings to detect single vortex rings. The total momentum associated with a ring is

$$\iiint \frac{\rho_s \hbar}{m_4} \mathbf{grad} S \, d^3 r = \iint \rho_s \kappa_0 \, d\mathbf{A} , \tag{2.12}$$

where the double integral is over an area bounded by the vortex ring. For a circular ring of radius R and a vortex core of radius a, this gives $\rho_s \kappa_0 \pi (R - a)^2$ in the direction normal to the plane of the ring. The expression for the speed of the ring is analogous to the expression for the magnetic field acting on a circular loop due to a current flowing round the loop, and is

$$v = \frac{\kappa_0}{4\pi R} \left(\ln \frac{8R}{a} - \frac{1}{4} \right) . \tag{2.13}$$

The expression for the energy is analogous to the expression for the magnetic energy of a current-carrying loop, and is

$$E = \frac{1}{2} \kappa_0^2 \rho_s R \left(\ln \frac{8R}{a} - \frac{7}{4} \right) + \text{const.} \tag{2.14}$$

The equations for energy, momentum and velocity are of the form that make the area of the ring and the position of the plane of the ring conjugate variables. This was exploited by Volovik [29] in his discussion of the quantum

tunneling of vortex rings. This Lagrangian formulation of vortex dynamics goes back more than a hundred years to Kirchhoff.

Rayfield and Reif worked at relatively low temperatures, around 0.5 K, and found that at such temperatures ions could flow without losing energy to the phonon system, apparently because they were trapped on vortex rings. The energies of the ions could be changed by known amounts by passing them through voltage drops, and the speed could be measured by a resonance technique that involved the coherent voltage pulses applied as they moved. The results show that the energy and speed are roughly reciprocals of one another, as larger energy means a larger vortex ring and so a lower speed. In terms of current–voltage relations this means that the product of current and voltage is roughly constant. The quantity obtained most directly is

$$vE = \frac{\kappa_0^3 \rho_s}{8\pi} \left(\ln \frac{8R}{a} - \frac{7}{4} \right) \left(\ln \frac{8R}{a} - \frac{1}{4} \right), \qquad (2.15)$$

but more careful fitting gives the best value of the vortex core radius a (apparently somewhat less than the interatomic spacing at those temperatures), and so gives $\kappa_0^3 \rho_s$. The precision with which κ_0 is determined in these experiments is comparable with the precision of the Vinen experiments.

Although it is much easier to detect magnetic flux than circulating superfluid, direct measurements of quantized flux in superconductors were not particularly easy. Two experimental measurements were published in 1961, by Deaver and Fairbank [30] and by Doll and Näbauer [31]. These had an accuracy of about 20%. A short time later, measurements of the fluxoid, as given in equation (2.9), were made by Parks and Little [32]. Modern measurements are somewhat more precise, and a measurement of the flux quantum for copper oxide superconductors by Gough et al. [33] showed that changes in flux were quantized to a value of $h/2e$ with a precision of about 4%. One reason for the poor precision is that a direct measurement of flux usually depends on a detailed knowledge of the geometry of the sample, and the position of the magnetometer.

Abrikosov's [34] prediction that the magnetic flux should penetrate a Type II superconductor as a regular lattice of flux lines was first verified by Cribier et al. [35] using neutron diffraction. Essmann and Träuble [36] developed a technique of decorating the regions of strong magnetic field with magnetic particles to show the flux lattice, which usually has dislocations and other defects, directly. These measurements were also used to compare the flux density with the number of lattice points per unit area, to confirm the magnitude of the flux quantum in this context. Recent work on the copper oxide superconductors using neutron scattering can be found in the work of Cubitt et al. [37], and using the decoration method can be found in the work of Bishop et al. [38]. In these materials the very large anisotropy

between directions normal to and parallel to the copper oxide planes is important, and vortices that are well localized to lattice points within a plane may wander between planes.

The flux lattice is extremely important for understanding the electrical resistance of Type II superconductors in the mixed state. In a nearly perfect material the flux lattice would flow with the electric current, just as vortices in superfluid ^{4}He are carried along by the superfluid flow. Irregularities would then cause the vortices to lower their energy by drifting transversely to the current direction, and such transverse motion can be shown to generate a voltage in the direction of the current, so that the motion is dissipative. Increased disorder pins the vortex lattice, so that the current can flow through it without dissipation. Thus the flux flow resistivity is decreased by increasing disorder, and good materials for superconducting magnets are highly disordered.

Use of the decoration method for detecting vortex lines in rotating superfluids, and took many years of work by by Packard *et al.* [39, 40]. The method involved trapping ions on the equilibrium (or steady state) vortices of a rotating cylinder of superfluid, and then ejecting the ions to get a photographic image of the positions of the vortices. The structure of neutral superfluid rotation is much more difficult to stabilize and to display than the structure of magnetized superconductors.

2.3   Precision of circulation and flux quantization measurements

Theoretical arguments suggest that the only limit to the precision with which flux is quantized in a ring of superconductor below the critical field (B_{c1} in the case of a Type II superconductor) is set by the magnitude of the term $\oint \mathbf{j} \cdot d\mathbf{R}$ on the right side of equation (2.9). Since the current density is governed by the London equation (2.7), it becomes exponentially small in the interior of a sample large compared with the penetration depth.

The Josephson effects [41–43] depend essentially on the quantization of flux. The SQUID magnetometer gives a response which is periodic in the fluxoid, which is almost equal to the flux except for a small contribution from the current density in the neighborhood of the weak link itself. I am not aware of any very precise absolute calibration of a SQUID, or of precise comparison between SQUIDs made of different materials. A less direct application of flux quantization is provided by the use of the ac Josephson effect to measure voltages by the relation

$$V = \frac{h}{2e}\nu \qquad (2.16)$$

between voltage and frequency. The connection of this with flux quantization is that the emf (electromotive force) round a circuit can be written

as

$$\oint \mathbf{E} \cdot d\mathbf{r} = -(d/dt) \oint \mathbf{A} \cdot d\mathbf{r} \tag{2.17}$$

in a gauge in which the scalar potential vanishes, which is the rate of change of flux, as is immediately apparent from Faraday's law. The equation gives a connection between the number of flux quanta per unit time as given by Faraday's law and the ac frequency generated. This has been used for high precision measurements. The Josephson voltage standard is the best voltage standard that there is, whose adoption led to revisions of the accepted values of fundamental constants [44,45]. Voltage balance between Josephson junctions made from different materials have shown a relative precision of a few parts in 10^{17} [46–48].

In a neutral system the circulation of the superfluid velocity is a topological quantum number, and is therefore exact in principle. However, there is a sense in which this is tautologous, at nonzero temperatures or in thin films, since the superfluid velocity is *defined* as the gradient of a phase angle. Physically measurable variables are the total fluid density, the average mass flow, and the normal fluid velocity, which is set by the physical boundaries. Superfluid density is determined by combining these variables together, for example by equating the mass flow to $\rho_n \mathbf{v}_n + \rho_s \mathbf{v}_s$. In the Vinen experiment it is not the circulation of the superfluid velocity itself which is measured, but, if equation (2.10) is accepted, it is the circulation of mass flow (momentum density). Since this question is the focus of much of our current work, it is discussed separately in Section 3. Since the vortex ring experiments also depend on the energy–momentum relations of vortex rings, I think that they are dependent on the same sort of relation.

Even at low temperatures in a bulk system it is not clear with what precision this type of experiment could be used to determine it even if the experimental difficulties could all be overcome. The results seem to depend not only on the quantized circulation, but also on rather specific details of the two-fluid dynamics. It is not a fundamental problem that the superfluid density is needed in equation (2.11), since this can be measured independently, and, for superfluid ^{4}He, it rapidly approaches the mass density at low temperatures. A more serious problem is that the frequency shift is proportional to the ratio of the superfluid density to the effective mass per unit length of the wire, and this effective mass includes the hydrodynamic mass of the surrounding fluid. This hydrodynamic mass is not simply the mass of fluid displaced by the wire, as it is for an ideal incompressible classical fluid, but there are uncertain corrections due to boundary layer effects, and, as has been pointed out by Duan [49] and Demircan et al. [50], there is a correction due to the compressibility of the fluid that is logarithmically divergent for low frequencies and large systems. Since helium has a much

lower density than the wire this is not a major correction, but it is one factor that makes it difficult for the experiment to be very precise.

Another possible source of imprecision is due to the effects of the boundary on the flow in the interior. Under ideal geometrical conditions, with a very small wire at the axis of a cylinder of radius R, displacement of the wire by a small distance a from the axis of the cylinder produces an effect that can be represented by an image of the vortex at a distance R^2/a, so there is a backflow of magnitude $\kappa_0 a/2\pi R^2$, which leads to a correction of the Magnus force on the wire whose relative magnitude is hT/m_4R^2, where T is the period of oscillation of the wire. For the conditions of Vinen's experiment this gives a correction of a few parts in 10^5. Such effects are due to the fact that the controling equation for S is the Laplace equation in a neutral system, whereas for a superconductor long range effects are exponentially reduced for distances larger than the penetration depth.

3   The Magnus force

3.1   Magnus force and two-fluid model

This section is largely based on work that my collaborators and I have done in recent years. There is a brief review of this work contained in a paper written by five of us in the autumn of 1997 [51].

Before discussing the details of the theory I want to give a brief review of the theory of the Magnus force in classical hydrodynamics, and then discuss the modifications which may be needed as a result of the two-fluid picture of superfluidity.

It is an old paradox of classical hydrodynamics that potential flow around an object gives no drag in the direction of the fluid flow and no lift perpendicular to it. Drag is provided by the effects of viscosity and by the creation of turbulence, and is very complicated, but lift is produced by the interaction of circulation of the fluid round the object with its motion, and has, to lowest order, a very simple form. A partial explanation is given in many textbooks of elementary physics, and is applied to problems like the lift on the wing of an airplane or the curved trajectory of a spinning ball. The usual explanation is given in terms of the different Bernoulli pressures on the two sides of the object. Actually a rather more detailed explanation is needed, and a detailed explanation shows that the result is very general, and quite independent of details of the fluid such as whether it is compressible or incompressible.

Consider a cylinder, perhaps a solid cylinder, or perhaps the hollow core of a vortex, with circulation κ around it, held in a fixed position with fluid flowing past it with asymptotic velocity v_0 in the x direction. At a large distance R from the cylinder the components of velocity will be $(v_0 - \kappa\sin\theta/2\pi R, \kappa\cos\theta/2\pi R, 0)$. This gives a Bernoulli pressure which is

approximately

$$-\rho v_0^2/2 + \rho v_0 \kappa \sin\theta/2\pi R \,, \tag{3.1}$$

which gives a net force per unit length $-\rho v_0 \kappa/2$ in the y direction, acting on the fluid inside a cylinder of radius R. The total force acting on the cylindrical volume of fluid is this Bernoulli force plus the force which must be applied to keep the cylinder or vortex stationary, and this total force must equal the rate of change of momentum of the fluid which is instantaneously in the cylinder of radius R. The downward momentum density $\rho \kappa \cos\theta/2\pi R$ on the left side of the cylinder is replaced by upward momentum density on the right side as the cylinder of fluid moves from left to right with speed v_0, so the rate of change of momentum per unit length is

$$\frac{1}{L}\frac{\mathrm{d}P}{\mathrm{d}t} = \int_0^{2\pi} \frac{\rho\kappa v_0 \cos^2\theta}{2\pi R}\,\mathrm{d}\theta = \rho\kappa v_0/2 \,. \tag{3.2}$$

Comparison of this with the Bernoulli force shows that an additional force $\rho\kappa v_0$ must be applied in the y direction, or, alternatively, that the moving fluid exerts force $-\rho\kappa v_0$ on the vortex. The Galilean invariant form of this is

$$\mathbf{F}_\mathrm{M} = \rho\kappa\hat{\mathbf{k}} \times (\mathbf{v}_\mathrm{V} - \mathbf{v}_0) \,. \tag{3.3}$$

The argument depends only on the asymptotic properties of the flow, and on momentum conservation.

Despite its generality, this argument cannot be directly taken over to the case of a superfluid. A superfluid is described, both hydrodynamically and thermodynamically, by the two-fluid picture of Landau and Tisza. The phase of the condensate wave function determines the superfluid velocity $\mathbf{v}_\mathrm{s}$ through equation (2.1). At nonzero temperature there will be excitations from the condensate, phonons with a linear energy–momentum relation, and, in the case of ^{4}He, rotons with a nonzero wavenumber of the order of the reciprocal of the interatomic spacing, and an energy around 8 K. These excitations interact with one another to form a local equilibrium, and all the entropy of the system is concentrated in this normal fluid component. The spectrum is determined by the local value of $\mathbf{v}_\mathrm{s}$, but the average velocity is not, but is determined by the boundary conditions. In particular, there can be an equilibrium state in which the normal fluid velocity $\mathbf{v}_\mathrm{n}$ is zero because the boundaries are static, even when the superfluid velocity is nonzero. This was actually the situation in the experiment of Vinen [2] described in Section 2.2, where, in equilibrium (before the wire was made to vibrate) the normal fluid was at rest, but the superfluid was circulating around the wire down the axis of the cylinder.

The velocities $\mathbf{v}_\mathrm{s}$ and $\mathbf{v}_\mathrm{n}$ are essentially deduced from the boundary conditions rather than being directly measured. Since $\mathbf{v}_\mathrm{s}$ is defined as the gradient of a phase it does not make much sense to ask if its integral round

a closed loop deviates slightly from an integral multiple of h/m_4. Quantities that can be directly measured are the total fluid density

$$\rho = \rho_s + \rho_n \ , \tag{3.4}$$

and the total momentum density (mass flow)

$$\mathbf{p} = \rho_s \mathbf{v}_s + \rho_n \mathbf{v}_n \ , \tag{3.5}$$

and quantities such as the free energy density

$$F = F_0 + \rho_s \mathbf{v}_s^2/2 + \rho_n \mathbf{v}_n^2/2 \ , \tag{3.6}$$

and the entropy flow. From measurements of such quantities the variation of ρ_s and ρ_n with temperature and pressure can be deduced, and then the equation of state can be used to analyse other measurements.

The essential feature of two-fluid hydrodynamics that should be used to generalize equation (3.3) to the superfluid case is that the two fluids coexist without interfering with one another, as is shown by equation (3.6). Therefore if the superfluid and normal fluid circulations round a vortex are κ_s and κ_n, while the asymptotic superfluid and normal fluid velocities are $\mathbf{v}_s$ and $\mathbf{v}_n$, the two components will contribute independently to the transverse force, and give the result quoted from Putterman's book [23] in equation (2.10). Furthermore, the theory suggests that the superfluid circulation κ_s should be quantized, and the normal fluid circulation κ_n should not be stable, but should eventually be dissipated by the normal fluid viscosity.

3.2   Vortex moving in a neutral superfluid

The Magnus force itself provides some interesting connections between quantized variables. In classical mechanics such a nondissipative force linear in the velocity can be represented by a term in the Lagrangian which is linear both in the velocity and in the displacement. There is a lot of ambiguity in the definition of such a Lagrangian, since any total derivative of the form $\dot{\mathbf{r}} \cdot \nabla f(\mathbf{r}, t) + \partial f(\mathbf{r}, t)/\partial t$ can be added to it, but there is no ambiguity in the action round a closed path. This ambiguity is very similar to that introduced by a choice of gauge in electrodynamics. In quantum mechanics such a term in the Lagrangian translates into a Berry phase [52], a phase that depends on the path of the system but not on the speed with which the path is traversed – again, this phase depends on a choice of gauge, but the phase associated with a closed path is gauge independent. It was observed by Haldane and Wu [53] that the Berry phase associated with a vortex in a two-dimensional superfluid is an integer multiple of 2π when the vortex is taken on a closed path that surrounds an integer number of atoms.

This can be seen from equation (2.11), which gives a force derivable from a Lagrangian of the form

$$\mathcal{L} = \frac{1}{2}\rho_s\kappa_0(\mathbf{v}_s - \mathbf{v}_L) \times \mathbf{r}_L \,, \tag{3.7}$$

where ρ_s is now the superfluid density per unit area. Integration of this round a closed loop in the superfluid (where I take for simplicity $\mathbf{v}_s = 0$) gives an action equal to $\rho_s\kappa_0$ times the area of the loop, or, using $\kappa_0 = h/m_4$, h times the number of atoms enclosed, rescaled by the factor ρ_s/ρ. Since the Berry phase is the action divided by $\hbar$ this gives a Berry phase equal to 2π times the number of atoms enclosed. There is a similar relation between the Lorentz force on an electron in a magnetic field and the magnetic flux quantum h/e. The Berry phase associated with taking an electron on a closed path is equal to 2π times the number of flux quanta which the path surrounds.

These connections between different quantum numbers, circulation and number of atoms in one case, electric charge and magnetic flux in the other, have led my collaborators, Ao and Niu, and me to look more closely into the question [54]. The two-dimensional result of Haldane and Wu has an obvious generalization to three dimensions, where the vortex is defined not just by a point in two-dimensional space, but as a curve in three-dimensional space. When such a curve is moved around and then returned to its original position it sweeps out a two-dimensional surface, and the Berry phase should be equal to 2π times the number of atoms surrounded by this surface. This simple statement hides a number of difficulties that we have tried to address. We think we understand what is meant by the path of an electron, but what is meant by the path of a vortex, an object whose microscopic definition is obscure? The number of atoms inside a geometrical surface is not fixed, but is a quantity subject to zero-point as well as thermal fluctuations, so what number should be used in this context? The density of the superfluid is reduced at the vortex core by the Bernoulli pressure, so does this reduction in density reduce the Berry phase?

In a more recent paper, Thouless et $al.$ [55] have tried to sharpen some of the questions by considering the effect of pinning the vortex core to a certain curve $x_0(z,t), y_0(z,t)$ by centering a short-ranged potential (repulsive to the atoms) on this curve – the reduced density at the core should cause the core to be attracted to the curve. This has enabled us to study the dynamics of the vortex cleanly by studying the effects of moving the pinning potential. To determine the coefficient of v_V, we consider an infinite system with superfluid and normal fluid asymptotically at rest ($\mathbf{v}_n = 0 = \mathbf{v}_s$) in the presence of a single vortex which is constrained to move by moving the pinning potential. For simplicity we describe the two-dimensional problem of a vortex in a superfluid film, but the three-dimensional generalization is straightforward. Also we restrict this discussion to the ground

state of the vortex, but the generalization to a thermal equilibrium state is straightforward. The reaction force on the pinning potential is calculated to lowest order in the vortex velocity $\mathbf{v}_V$. This can be studied as a time-dependent perturbation problem, but this can be transformed into a steady state problem, with the perturbation due to motion of the vortex written as $i\mathbf{v}_V \cdot \mathbf{grad}_0$. The force in the y direction on a vortex moving with speed v_V in the x direction can then be written as

$$F_y = iv_V \left\langle \Psi_{\mathbf{r}_0} \left| \frac{\partial V}{\partial x_0} \frac{\mathcal{P}}{E_0 - H} \frac{\partial}{\partial y_0} \right| \Psi_{\mathbf{r}_0} \right\rangle + \text{ comp. conj.,} \qquad (3.8)$$

where $\mathcal{P}$ projects off the ground state of the vortex. Since $\partial V/\partial x_0$ is the commutator of H with the partial derivative $\partial/\partial x_0$, the denominator cancels with the H in the denominator, and so the expression is equal to the Berry phase form

$$F_y = -iv_V \left\langle \frac{\partial \Psi}{\partial x_0} \Big| \frac{\partial \Psi}{\partial y_0} \right\rangle + iv_V \left\langle \frac{\partial \Psi}{\partial y_0} \Big| \frac{\partial \Psi}{\partial x_0} \right\rangle . \qquad (3.9)$$

Since the Hamiltonian consists of kinetic energy, a translation invariant interaction between the particles of the system, and the interaction with the pinning center, which depends on the difference between the pinning center coordinates and the particle coordinates, the derivatives $\partial/\partial x_0$, $\partial/\partial y_0$, can be replaced by the total particle momentum operators $-\sum \partial/\partial x_j$, $-\sum \partial/\partial y_j$. This gives the force as a commutator of components, P_x, P_y of the total momentum,

$$F_y = -iv_V \langle \Psi_{\mathbf{r}_0} | [P_x, P_y] | \Psi_{\mathbf{r}_0} \rangle . \qquad (3.10)$$

At first sight one might think that the two different components of momentum commute, but this depends on boundary conditions, since the momentum operators are differential operators. Actually this expression is the integral of a curl, and can be evaluated by Stokes' theorem to get

$$F_y = v_V \oint \left\langle \Psi_{\mathbf{r}_0} \left| -i \sum_j \mathbf{grad}_j \right| \Psi_{\mathbf{r}_0} \right\rangle \cdot d\mathbf{r} = \oint \mathbf{j}_s \cdot d\mathbf{r}, \qquad (3.11)$$

where the integral is taken over a loop at a large distance from the vortex core. This gives the force in terms of the circulation of momentum density (mass current density) at large distances from the vortex.

There is actually a striking resemblance between the expression (3.10) for the coefficient of the Magnus force and the expression for the Hall conductance in terms of a Chern number which is discussed in Section 4.4.

Our result, that the transverse force is equal to v_V times the line integral of the mass current, is independent of the nature or size of the pinning

potential. The general form of this is

$$\mathbf{F}_t = \rho_s \mathbf{K}_s \times \mathbf{v}_V + \rho_n \mathbf{K}_n \times \mathbf{v}_V \,, \tag{3.12}$$

where $\mathbf{K}_n$ represents the normal fluid circulation, in agreement with the v_V-dependent part of equation (2.10).

Since our result is controversial, Tang and I have checked the method that we use by using it also to calculate the dissipative part of the force, the component in the direction of motion [56]. Although we are not able to evaluate this in the general case, we find that for symmetrical potentials acting on noninteracting particles the method gives the usual expression for the longitudinal force on a moving potential in terms of the transport cross section.

By using this technique we verify part of equation (2.10) without using any specific model of a vortex, independently of whether its core is a solid cylinder or the consequence of a mathematical singularity in the pure fluid. We find that the coefficient of $\kappa_0 v_V$ is indeed ρ_s, the asymptotic value of the superfluid density at large distances from the vortex core. Only the coherent part of the wave function contributes to the Magnus force, so at nonzero temperatures the Berry phase is reduced by a factor ρ_s/ρ.

The other half of equation (2.11), which gives the coefficient of $\kappa_0 v_s$, has been verified in recent work by Wexler [57]. Wexler considers superfluid contained in a ring, scuh as the one shown in Figure 2 with n quanta of circulation trapped in the ring, so that the superfluid velocity is $v_s = hn/m_4 L$, where L is the perimeter of the ring. If an additional vortex is greated on the outer edge of the ring, dragged slowly across the ring by a pinning force, and annihilated on the inner edge, the number of trapped quanta of cirulation is increased to $n + 1$, changing the superfluid velocity by $h/m_4 L$. This leads to an increase in the free energy of the superfluid circulating aroung the ring by

$$\Delta F = \rho_s \frac{h v_s}{m_4 L} LA = \rho_s \kappa_0 v_s A \,, \tag{3.13}$$

where A is the area of cross section of the ring, since the superfluid density ρ_s is defined in terms of the dependence of the free energy on superfluid velocity. The work done against the transverse force in moving a vortex across the ring is equal to the force per unit length times the area of cross-section A. Comparison of this with equation (3.13) shows that the transverse force per unit length acting on a nearly stationary vortex when superfluid flows past it with velocity v_s has magnitude $\rho_s \kappa_0 v_s$, in agreement with equation (2.10).

This argument has considerable analogies with the argument given by Laughlin for the integer quantum Hall effect [58], which is discussed in Section 4.3. In that argument the flux through a ring is changed, and electrons are moved from one edge of the ring to the other. Comparison is

made between the work done by the transverse force on the flux produced by the electric current and the energy change of the electrons moving from one edge to the other.

The surprising feature of these results is that the normal fluid component does not seem to enter into the expression for the transverse force on a vortex. If the coefficients of $\kappa_s v_V$ and $\kappa_s v_s$ are respectively $\pm\rho_s$, Galilean invariance shows that the coefficient of $\kappa_s v_n$ for the transverse force must be zero. Volovik [59] argues that we are wrong, and the coefficient of $\kappa_s v_V$ should be ρ rather than ρ_s, which then gives the coefficient of $\kappa_s v_n$ as $-\rho_n$, again from Galilean invariance. Our result is in contradiction with widely accepted results, going back to the work of Iordanskii [60], which show a transverse force, proportional to the normal fluid velocity, due to the scattering of phonons or rotons from a vortex. Sonin [24] has given a recent survey of this argument.

The arguments developed by us suggest that the integral of momentum density round a closed loop is actually the topological quantum number that can be measured, rather than the integral of superfluid velocity. The superfluid density, the factor by which these two quantities differ, is, how-ever, a quantity that can itself vary, not only with temperature, but also with velocity, so the quantization of this quantity is not very precise.

3.3   Transverse force in superconductors

The problem of how to generalize to a superconductor our arguments about the transverse force on a moving vortex in a neutral superfluid raised some difficulties. Translation invariance plays an essential part in the result of Thouless, Ao and Niu [55], and there are various features of superconduc-tivity that make translation invariance difficult to apply. The regular lattice of ions even in a perfect metal has only discrete translation invariance, not the continuous translation invariance that is needed for this argument. The impurities that exist in any real metal, and are essential for getting a fi-nite conductivity in the normal state, destroy even the discrete translation invariance. Finally, magnetic effects, which are essential for understanding properties of superconductivity such as the Meissner effect, are usually put in by means of a vector potential, and a choice of gauge for this hides the fundamental translation invariance, even if it does not destroy it.

We generalized the argument for an unrealistic model in which impu-rities and disorder are ignored, and in which the background of a regular array of positive ions is replaced by a uniform positive background, and some short-ranged pairing interaction between the electrons is put in to give superconductivity at low temperatures [61]. The magnetic field is not put in explicitly, but, in addition to the Coulomb interaction between the electrons, and between the electrons and the background, the current–

current interaction is included in the form

$$H_{\text{mag}} = -\frac{e^2\mu_0}{16\pi m^2} \sum_{n \neq n'} p_n^i \left[\frac{\delta^{ij}}{|\mathbf{r}_n - \mathbf{r}_{n'}|} + \frac{(\mathbf{r}_n - \mathbf{r}_{n'})^i(\mathbf{r}_n - \mathbf{r}_{n'})^j}{|\mathbf{r}_n - \mathbf{r}_{n'}|^3}) \right] p_{n'}^j \; . \quad (3.14)$$

This gives a complete account of electromagnetic effects, apart from the spin-orbit interaction and the relativistic mass correction, up to second order in v^2/c^2. It is interesting to note that this approach was taken by Dirac *et al.* [62] in responding to criticism by Eddington of the relativistic theory of the hydrogen atom – of course they also included spin-orbit effects and the second order mass correction, since it was the Dirac equation which was in question. We can ignore the mass correction and spin-orbit interaction because they are not important for superconductivity theory. We have shown that such an approach to superconductivity theory, without explicit introduction of magnetic fields produced by the electrons, does lead to the Meissner effect and to magnetic screening with the usual penetration depth.

Just as we do for the neutral superfluid, we introduce a pinning potential to control the position of the flux line. This is the only part of the Hamiltonian that breaks translation symmetry. The argument for the force on the pinning potential goes through as before, and we get the same result that the transverse force on a moving vortex is equal to the vortex velocity times the asymptotic value of the circulation of momentum density.

The magnitude of this transverse force for a vortex moving in an idealized superconductor is not surprising. It is equal and opoposite to the transverse Lorentz force which is obtained when a supercurrent flows past a stationary vortex. The value of the Lorentz force due to a supercurrent can be derived by an adaptation of Wexler's argument for the force due to flow past a vortex in a neutral superfluid [57], but there has been no serious doubt of this result. In combination these results tell us that in the absence of other forces, such as pinning and frictional forces, the vortices will flow with the average velocity of the supercurrent.

The form of the result is surprising, since at large distances the circulation of canonical momentum density does not correspond to any current density, and there is no Bernoulli pressure imbalance or net momentum flux which could be used to explain this force, in the way that the classical Magnus force is explained. The explanation was given thirty years earlier in a paper by Nozières and Vinen [63]. At distances small compared with the penetration depth the moving magnetic flux line does behave like a vortex, and the forces are mostly hydrodynamic. The magnetic flux that moves with the line does, however, generate an electric field, a dipolar field, which exerts a net force on the rigid positive substrate. Therefore this transverse force, which was hydrodynamic close to the vortex core, is transmitted to large distances as an elastic stress in the positive background.

In fact the arguments and results of Nozières and Vinen are very close to ours. Like us they assumed that the positive background was uniform. We have not had to assume, as they had to, that the superconductor is very strongly Type II.

4 Quantum Hall effect

4.1 Introduction

The quantum Hall effect was discovered in 1980 by von Klitzing *et al.* [64]. Systems which display the quantum Hall effect are generally two-dimensional electron systems such as are found at the interface between silicon and silicon dioxide in an MOS (metal-oxide-semiconductor) device, or at the interface of a heterojunction such as the GaAs-Al$_x$Ga$_{1-x}$As system. A strong magnetic field (5 T or more) and low temperatures (helium temperatures or less) are needed to observe the effect.

One of the striking things about this initial report is that quantization of the Hall conductance I/V_H was observed to be an integer multiple of e^2/h with very high precision, better than one part in 10^5. Later measurements have shown an absolute precision of 1 part in 10^7 [65]. A comparison in which the Hall voltage generated in a silicon MOS device against the Hall voltage generated by the same current in a GaAs-Al$_x$Ga$_{1-x}$As device has shown consistency between different devices of a few parts in 10^{10} [66]. The quantum Hall effect is sufficiently reproducible that it provides the best available secondary standard of electrical resistance, and its value is included in the adjustments of fundamental constants [67]. This precision suggests that a topological explanation of the quantum Hall effect is appropriate.

The earliest theoretical interpretations of the observed quantization showed that the plateaus in the Hall conductivity came from filled Landau levels, and that the Fermi level was pinned between Landau levels by localized states produced by the disorder of the substrate; these localized states make no contribution to the low-temperature conductivity. The quantized value is unaltered by disorder and interactions to all orders in perturbation theory [68–70]. These arguments, although sound, do not connect the quantum Hall effect with other phenomena that have a very high precision, but Laughlin [58] gave an argument which is much more general, and which revealed a topological basis for the integer quantization. Later work [71–75] has interpreted the topological aspects of the effect differently, but Laughlin's argument remains one of the most powerful ways of understanding the quantum Hall effect.

As soon as the reasons for integer quantization of the Hall conductance were understood clearly, experiments by Tsui *et al.* [76] showed that the Hall conductance could be a fractional multiple of e^2/h. The initial work was not very precise, but later the fractional values were shown very clearly as

plateaus in the Hall conductance, whose precision was less well determined than that of the integer effect only because lower temperatures and higher magnetic fields were needed.

4.2   Proportionality of current density and electric field

The proportionality of electric field and electric current density perpendicular to the field for a two-dimensional electron system can be explained in a simple and straightforward manner, and should have led to a prediction of the quantum Hall effect before it was discovered experimentally. For a system of noninteracting electrons in the absence of a substrate potential the Hamiltonian can be written as

$$
\begin{aligned}
H &= \frac{1}{2m}\left(-i\hbar\frac{\partial}{\partial x}+eA_x\right)^2 + \frac{1}{2m}\left(-i\hbar\frac{\partial}{\partial y}+eA_y\right)^2 + V(x,y) \\
&= -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{1}{2m}(\hbar k_y + eBx)^2 + e\mathcal{E}x \,,
\end{aligned}
\tag{4.1}
$$

where the Landau gauge $A_x = 0$, $A_y = Bx$ has been used, the electric field is $\mathcal{E}$ in the x-direction, and the y-dependence of the wave function is taken to be $\exp(ik_y y)$. The y-component of the current density operator is $e(\hbar k_y + eBx)$, and so this Hamiltonian leads to a current density corresponding to an electron drift velocity $-\mathcal{E}/B$ in the y-direction, just as is found in the classical theory of charged particle motion. For n full Landau levels the electron density is nBe/h, n times the density of single electron flux quanta, so the current density is

$$
j_y = (nBe/h)(e\mathcal{E}/B) = (ne^2/h)\mathcal{E} \,.
\tag{4.2}
$$

This gives the right result, but is not yet an adequate explanation of the observed quantization of the Hall conductivity σ_{xy}, because it does not give plateaus as the magnetic field is varied for fixed electron density, since there is a partially filled Landau level for general values of the magnetic field B. Under some conditions, such as in silicon MOS devices, it is the chemical potential rather than the electron density that is controlled, but in other systems, such as optically excited GaAs systems, the recombination time is very long, and the electron number is kept constant.

A refinement of this argument [68–70] shows that weak disorder makes no change in this result, to all orders in perturbation theory. The disorder produces localized states, lying between the mobile states associated with each Landau level, and these localized states serve to pin the Fermi energy between Landau levels without changing the Hall current associated with the mobile states.

The Kubo formula [77, 78], which relates the conductivity of a material to its current-current correlation function, can be used to display this result

in a form in which its insensitivity to perturbations is made manifest. For the longitudinal conductivity in the dc limit the Kubo formula involves a delta-function on the Fermi surface, and is rather sensitive to the order of limits, but the result for the Hall conductivity is less sensitive. In a many-electron state $|\Psi_0\rangle$ of energy E in which no current flows, the electron current in the y-direction induced by the perturbation $e\mathcal{E}\hat{X}$ is, to lowest order in perturbation theory,

$$\mathcal{E}e\langle\Psi_0|\hat{J}_y\frac{\mathcal{P}}{E-\hat{H}}\hat{X}|\Psi_0\rangle + \mathcal{E}e\langle\Psi_0|\hat{X}\frac{\mathcal{P}}{E-\hat{H}}\hat{J}_y|\Psi_0\rangle\,, \qquad (4.3)$$

where $\mathcal{P}$ is the operator that projects off the eigenstate $|\Psi_0\rangle$ of $\hat{H}$. Since the current operator $\hat{J}_x$ is given by the commutation relation $[\hat{H},\hat{X}]=-i\hbar\hat{J}_x/e$, this gives the current in the form

$$
\begin{aligned}
J_y &= -i\hbar\mathcal{E}\langle\Psi_0|\left[\hat{J}_y\frac{\mathcal{P}}{(E-\hat{H})^2}\hat{J}_x - \hat{J}_x\frac{\mathcal{P}}{(E-\hat{H})^2}\hat{J}_y\right]|\Psi_0\rangle \\
&= \frac{\hbar\mathcal{E}}{2\pi}\oint dz\,\mathrm{Tr}\frac{1}{z-\hat{H}}\hat{J}_y\frac{1}{z-\hat{H}}\hat{J}_x\frac{1}{z-\hat{H}}\,, \qquad (4.4)
\end{aligned}
$$

where the integral over the complex variable z goes round only the lowest of the eigenvalues of the Hamiltonian $\hat{H}$, the ground state and neighboring states that involve localized excitations from the ground state. The integrand is closely related to a Green function at energy z, and such a Green function is exponentially localized both when the imaginary part of z is nonzero, and when z lies in a gap in the spectrum or in a region of the spectrum where eigenstates are localized by disorder [79] (in a mobility gap). This argument, or a simple modification of it, shows that the result is unchanged by local perturbations that are not strong enough to push extended states through the Fermi surface, and that the current density is a local function of the applied field when the Fermi energy lies in a mobility gap.

4.3 Bloch's theorem and the Laughlin argument

In the early 1930s Bloch proved a result that he claimed showed that all existing theories of superconductivity were wrong – this was just before the Meissner effect was discovered. Bloch's theorem states that the free energy F of the equilibrium state for a loop or other nonsimply connected piece of conductor is a periodic function of the flux Φ enclosed by the loop, with period h/e, so that the current $\partial F/\partial\Phi$ is periodic and has zero average. This result was widely known, but was never published by Bloch, and one of the best discussions was given by Bohm in 1949 [80]. The oscillations about zero of the equilibrium current are generally expected to be small,

but the example of the SQUID magnetometer shows that such oscillatory equilibrium currents can be relatively large under certain circumstances.

Laughlin's argument for the quantization of Hall conductance is compact and powerful, but perhaps too compact to be readily understood. A more transparent version of Laughlin's argument was given by Halperin [81]. I have rephrased Laughlin's argument as a generalization of Bohm's version of the Bloch theorem [82]. I consider an annulus of two-dimensional conductor in a uniform magnetic field, with a strength such that the Fermi energy lies in a mobility gap. The Bloch theorem tells us that in equilibrium the only current flowing is the mesoscopic current which averages to zero when the flux threading the annulus is varied. In Laughlin's argument the electrochemical potentials at the two edges of the annulus are allowed to have slightly different values, μ_i on the inner edge and μ_o on the outer edge. Because there are no mobile electron states with energies close to the range between μ_i and μ_o, this nonequilibrium state can be maintained. I suppose that the electrochemical potentials are defined by keeping reservoirs of electrons with Fermi energies μ_i and μ_o in contact with the two edges. If the flux Φ threading the annulus is now changed by one quantum unit $\delta\Phi = h/e$ the annulus returns to its former state, apart from trivial gauge changes of the electron states by the factor

$$\exp\left[-i(e/\hbar)\int \delta\mathbf{A}(\mathbf{r})\cdot\mathrm{d}\mathbf{r}\right] . \tag{4.5}$$

Such a gauge change is allowed even for electron states that extend round the annulus, since the wavefunction remains single valued. Since the annulus has returned to its original state, apart from this gauge change, the only significant thing that could have happened is that an integer number n of electrons might have passed across the system from the inner reservoir to the outer. The change in free energy of the system is therefore $n(\mu_o - \mu_i)$. The work done is the time integral of the current times the voltage around the annulus, and this voltage is $\mathrm{d}\Phi/\mathrm{d}t$, by Faraday's law, so that the equality of these two gives

$$\int J\frac{\mathrm{d}\Phi}{\mathrm{d}t}\,\mathrm{d}t = \int J\,\mathrm{d}\Phi = n(\mu_o - \mu_i) . \tag{4.6}$$

The left hand side is $\bar{J}(h/e)$, where $\bar{J}$ is the current round the ring, averaged over the fractional part of the flux, and the right side is neV, where $V = (\mu_o - \mu_i)/e$ is the voltage between the two edges, so the conclusion of Laughlin's argument is that

$$\bar{J} = n\frac{e^2}{h}V , \tag{4.7}$$

and the conductance is quantized as an integer multiple of e^2/h.

At first sight it is surprising that electrons can be moved across the system in the absence of any mobile states close to the Fermi energy, but a more detailed explanation is given in Halperin's paper [81]. Each occupied Landau level has edge states that are close to the Fermi energy, and there is a quasicontinuum of mobile states between the two edges, with the states in the interior well below the Fermi energy. The effect of changing the vector potential is to make an adiabatic change that maps this continuum of states into one another in such a way that n states on the inner edge are emptied and n states on the outer edge are filled. For example, if we take the case of an ideal annulus and use the symmetric gauge centered on the center of the annulus, the states of the lowest Landau level can be written in the form

$$\psi(x,y) = |x+iy|^{-e\Phi/h}(x+iy)^s \exp\left(-|x+iy|^2\frac{eB}{4\hbar}\right), \qquad (4.8)$$

where s is an integer. This wave function is concentrated around a circle of radius r_s, where

$$B\pi r_s^2 = -\Phi + sh/e ; \qquad (4.9)$$

this is a good wave function provided the value of r_s lies between the inner and outer radii of the annulus. Decrease of Φ by h/e pushes each wave function out to the former position of the next one, and so moves one electron from the inner edge to the outer edge. A similar argument holds for higher occupied Landau levels.

This argument is very general, and all it seems to require is that the Fermi energy lies in a mobility gap. It does not, for example, require that the voltage difference between the two edges should be vanishingly small. All that is required is that the voltage difference should be small enough that there is no appreciable tunneling of electrons between the two edges or to higher unoccupied levels. The only unsatisfactory feature of the argument is that it gives only the current $\bar{J}$ averaged over the flux in the annulus, whereas the actual current could include a mesoscopic contribution. Such a mesoscopic contribution could only come from the edge states, since the bulk Green functions are exponentially localized at the Fermi energy. In the edges of typical devices used for studying the quantum Hall effect there are many levels contributing to edge currents and to diamagnetic susceptibility, and there is no reason to expect strong interference effects that could give rise to a significant mesoscopic correction to equation (4.7).

Because one does not expect corrections to be propagated over long distances, one also expects that the details of the geometry should not affect the current-voltage relation. The same relation that Laughlin showed for the annulus should also hold for a more typical arrangement where current is fed into a Hall bar at one end and removed from it at the other.

An important aspect of this argument is that it suggests sources for departures from exact quantization, as well as the order of magnitudes

to be expected for them. Firstly, there are processes that lead to currents between the edges as well as round the annulus, and these can alter the estimates of free energy changes used in equation (4.6). These longitudinal currents can be produced by thermally activated hopping of localized electrons, thermal activation of electrons in unfilled Landau levels or of holes in filled Landau levels, or by quantum tunneling. The last of these will fall exponentially as the width of the annulus increases, and the others fall exponentially as the temperature is lowered, so these effects can be made exponentially small by increasing the width of the system and reducing the temperature. The other obvious source of deviations from exact quantization is that the actual current may deviate from its value $\bar{J}$ averaged over a flux period. There is no reason to expect such deviations to be large, as they are in a Josephson junction, and the simplest picture of them is that they come about as the closely spaced edge levels move across the Fermi energy under the influence of changing values of the flux Φ. The energy spacing between such edge levels should be inversely proportional to the circumferences of the edges, and might give rise to corrections inversely proportional to the circumference at zero temperature, but I would expect thermal broadening to reduce these corrections to something exponential once the temperature exceeds the mean energy spacing between edge states.

This argument contains the essential features as the more explicitly topological argument which is presented in the next subsection. It uses the gauge covariance of quantum mechanics in regions that are not simply connected, together with the quantization of electron charge in the reservoirs. It is clearly invariant under small perturbations, so long as they do not bring the energies of mobile electrons in the interior close to the electrochemical potential range. It also has some features similar to the argument for the force on a vortex in a supercurrent which was presented in Section 3.2 [57]. A charge carrier in the theory of the quantum Hall effect plays much the same role as a vortex in a superfluid.

4.4   Chern numbers

The first expression for the Hall conductance which gave explicitly a topological invariant was obtained for the case of electrons moving simultaneously in a uniform magnetic field and a periodic potential [71]. This is a problem for which very interesting results had previously been obtained by Azbel [83] and Hofstadter [84]. A weak periodic potential splits each Landau level into q subbands, where there are q flux quanta for every p unit cells of the periodic potential. Each of these subbands carries an integer Hall conductance, and these integers can be different from unity. For example, when $p/q = 3/5$ the Hall conductances of the 5 subbands alternate between -1 and 2.
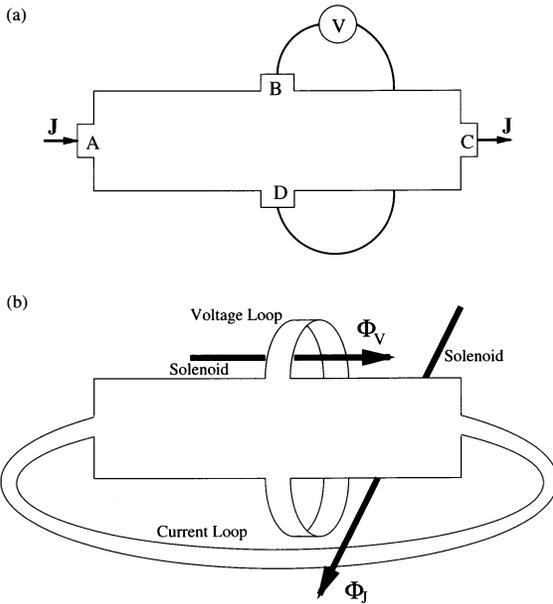
Fig. 3. The Hall bar, with current and voltage leads, shown in (a) can be replaced by the arrangement shown in (b), where the voltage is supplied by changing flux Φ_V through one loop, and the current is monitored by observing changes of the flux Φ_J through the other loop.

The quantum number characterizing the Hall conductance of a subband turns out to be the topological invariant known as the Chern number.

Later work showed that this method could be extended to a much more general situation. In order to observe the Hall effect in the usual way one takes a bar of the two-dimensional electron system in a strong magnetic field, passes a fixed current through it from a pair of current leads, and measures the voltage across the sample by connecting two voltage leads on opposite edges of the sample to a voltmeter. This is the set-up shown in idealized form in Figure 3a. In the work of Avron and Seiler [73], and of Niu and Thouless [85], the leads connected to current source and voltmeter are replaced by leads of the same material as the Hall bar, connected in pairs as shown in Figure 3b. Through the voltage loop there passes a solenoid which has a variable flux Φ_V, while there is another solenoid with flux Φ_J passing through the current loop. If the flux Φ_V is changed at a uniform rate, the solenoid will maintain a constant electromotive force $d\Phi_V/dt$ around the voltage loop, and the other solenoid can be used as a pick-up to monitor the current that is generated around the current loop. One may notice

that while experimentalists like to measure the voltage resulting from the passage of a fixed current, theorists prefer to ask about the current that will be generated by a given voltage. The current operators in the current and voltage leads are $\partial H/\partial \Phi_J$ and $\partial H/\partial \Phi_V$.

The Hall conductance can be calculated from the Kubo formula [77, 78] for the conductivity, which gives, by means of standard perturbation theory, the current density which is the linear response to an applied electric field or electrochemical potential gradient. The result can be written in terms of a current-current correlation function, as it was in equation (4.4). For the Hall conductance we want the J_x, J_y correlation function for the many-electron ground state wave function $|\Psi_0(\Phi_J, \Phi_V)\rangle$, and so the Hall conductance can be written as

$$
\begin{aligned}
S_H(\Phi_J, \Phi_V) \quad = \quad & i\hbar\langle\Psi_0(\Phi_J, \Phi_V)|\left(\frac{\partial H}{\partial \Phi_V}\frac{\mathcal{P}}{(E-H)^2}\frac{\partial H}{\partial \Phi_J}\right. \\
& \left. -\frac{\partial H}{\partial \Phi_J}\frac{\mathcal{P}}{(E-H)^2}\frac{\partial H}{\partial \Phi_V}\right)|\Psi_0(\Phi_J, \Phi_V)\rangle, \quad (4.10)
\end{aligned}
$$

where H is the hamiltonian for the system, depending on the parameters Φ_J and Φ_V, and E is the corresponding energy of the ground state. The operator $\mathcal{P}$ projects off the ground state. Perturbation theory for the wave function gives

$$
|\frac{\partial \Psi_0}{\partial \Phi_V}\rangle = \frac{\mathcal{P}}{E-H}\frac{\partial H}{\partial \Phi_V}|\Psi_0\rangle, \quad (4.11)
$$

and the corresponding equation for the perturbation due to the flux through the current loop, so the the equation for the Hall conductance is

$$
S_H(\Phi_J, \Phi_V) = i\hbar\left(\langle\frac{\partial \Psi_0}{\partial \Phi_V}|\frac{\partial \Psi_0}{\partial \Phi_J}\rangle - \langle\frac{\partial \Psi_0}{\partial \Phi_J}|\frac{\partial \Psi_0}{\partial \Phi_V}\rangle\right). \quad (4.12)
$$

This must be periodic in each of the fluxes with period h/e, and the fluxes have the effect of changing the phase of quasiperiodic boundary conditions round each of the current and voltage loops.

The quantities that appear in this equation can be written in terms of the Green function for the many-body system, integrated over the two spatial coordinates of the system, and integrated around a contour in the complex energy plane which surrounds a part of the real axis that includes no mobile excited states. Since the Fermi energy lies in a mobility gap, we expect the Green function to fall off exponentially with distance at a rate that depends on the localization length at the Fermi energy, except at the edge of the system where there will be extended electron edge states [81]. In this system the edge states cannot contribute to the total current around the current loop, since there is only one edge, and any current that flows along the edge goes in opposite directions on the two edges of the current leads.

The system with its two sets of leads looping around has the topology of a torus pierced by a single hole. I do not think that this argument has been properly worked out in detail, but it suggests that the expression (4.12) is independent of the quasiperiodic boundary conditions determined by Φ_J, Φ_V, up to corrections exponentially dependent on the ratio of the width (or length) of the Hall bar to the localization length at the Fermi energy. If this is the case we can write the Hall conductance as

$$S_H = \frac{ie^2}{2\pi h} \int_0^{h/e} d\Phi_J \int_0^{h/e} d\Phi_V \left(\langle \frac{\partial \Psi_0}{\partial \Phi_V} | \frac{\partial \Psi_0}{\partial \Phi_J} \rangle - \langle \frac{\partial \Psi_0}{\partial \Phi_J} | \frac{\partial \Psi_0}{\partial \Phi_V} \rangle \right). \quad (4.13)$$

The wave function $|\Psi_0(\Phi_J, \Phi_V)\rangle$ gives a mapping of the torus defined by Φ_J, Φ_V (physical quantities are periodic in these variables) onto the complex projective space of normalized wave functions with arbitrary phase. The integral is 2π times the integer invariant that defines the first Chern class of this mapping [86].

For a physicist, a more familiar way of getting the desired result is to argue that the integrand is the curl of the vector whose components in Φ_J, Φ_V space are

$$\frac{1}{2} \left(\langle \frac{\partial \Psi_0}{\partial \Phi_J} | \Psi_0 \rangle - \langle \Psi_0 | \frac{\partial \Psi_0}{\partial \Phi_J} \rangle, \langle \frac{\partial \Psi_0}{\partial \Phi_V} | \Psi_0 \rangle - \langle \Psi_0 | \frac{\partial \Psi_0}{\partial \Phi_V} \rangle \right), \quad (4.14)$$

which is $-i$ times the gradient of the phase η of $|\Psi_0\rangle$ in this flux space. This therefore gives

$$S_H = \frac{e^2}{2\pi h} \oint \mathbf{grad}_\Phi \eta \cdot d\Phi = \frac{ne^2}{h}, \quad (4.15)$$

where the integral is taken round the boundary of the two-dimensional integration in equation (4.13). The phase, which is very like a Berry phase [52], must be defined in some unambiguous way, such as by parallel transport, or by fixing the phase of the wave function to be zero at some point in the space of electron coordinates – but then ambiguities arise at those values of Φ_J, Φ_V which give a zero at this chosen point. However the phase is defined, it must return to the same value around the path in equation (4.15) up to a multiple of 2π. It is this winding number of the Berry phase that gives the integer n on the right side of the equation. Various implementations of the phase have been discussed by Thouless [87], Kohmoto [88], and by Arovas et al. [89].

Although the result can be reduced to the winding number of a Berry phase round the perimeter of a unit cell in two-dimensional flux space (Φ_J, Φ_V), yet the argument seems to be intrinsically two-dimensional, involving simultaneously what is happening in the current leads and in the voltage leads. The same could be said of the Laughlin argument [58], where

there needs to be a simultaneous consideration of the change in vector potential around the annulus and the transfer of electrons across the annulus.

Again there is some analogy between the two-form in equation (4.14) that gives the Hall conductance and the expectation value of a commutator in equation (3.10) that gives the transverse force on a vortex moving in a superfluid in Section 3.2.

4.5   Fractional quantum Hall effect

The arguments that have been presented in the last two subsections show that the Hall conductance is an integer multiple of e^2/h if the Fermi energy lies in a mobility gap, and if the ground state wave function is unique. The discovery of fractional values of the Hall conductance number by Tsui *et al.* [76] was therefore surprising. Subsequent work has shown that many different simple odd-denominator fractions occur, and that the fractional quantization is fairly precise. Simple modifications of the theory for non-interacting or weakly interacting electrons did not seem to give this effect, and Laughlin [90] argued that it must be due to the existence of a new sort of correlated many-electron ground state, and proposed the sort of ground state that should reduce the repulsive Coulomb energy of the electrons and display fractional quantization with odd-denominator fractions.

In the central gauge, with the vector potential equal to $\mathbf{A} = (-By/2, Bx/2)$, the degenerate many-body ground state wave function for N non-interacting electrons all in the lowest Landau level has the form

$$f(z_1, z_2, \ \ldots \ z_N) \exp\left(-\frac{1}{4l_0^2}\sum_{i=1}^{N}|z_i|^2\right), \tag{4.16}$$

where z_i represents $x_i + iy_i$, f is any multinomial antisymmetric in the variables, and l_0 is the magnetic length $\hbar/eB$. If f has the form

$$f = \prod_{i<j}^{N}(z_j - z_i), \tag{4.17}$$

the particles are concentrated in a disk whose area is approximately $2\pi Nl_0^2$, and this represents a fully occupied Landau level with a density, inside the disk, of one electron per flux quantum. Laughlin [90] suggested that the wave function obtained by setting the multinomial equal to

$$f(z_1, z_2, \ \ldots \ z_N) = \prod_{i<j}^{N}(z_j - z_i)^q, \tag{4.18}$$

with q an odd integer, would be particularly effective at keeping the electrons apart and so reducing the Coulomb energy. It gives a uniform one-electron

density within the disk, and has no pairs of electrons whose relative angular momentum is less than $q\hbar$. Haldane and Rezayi [91] have shown that the generalization of this wave function to spherical geometry gives the true and unambiguous ground state for a model interaction similar to the Coulomb interaction. It is also in very close agreement with finite system calculations for the Coulomb interaction.

For such a wave function, given by equations (4.16) and (4.18), the highest power of z_j in the multinomial is $(N-1)q$, so the area of the disc is increased by a factor of q, and the density of electrons is now $1/q$ per flux quantum. This ground state is separated from excited states by an energy gap for the creation of quasiparticle excitations. The lowest energy quasiparticles are presumably localized by the substrate disorder, so the filling does not have to be exactly $1/q$ for the Hall conductance to be determined by the properties of this ground state. This fractionally occupied Landau level will give a Hall conductance equal to e^2/qh whether or not there are additional localized quasiparticles present. Laughlin showed that there should be a hole-like quasiparticle with fractional charge e/q formed at the point z_0 by adding an extra flux line at that point; this can be done by multiplying the ground state wave function by a constant times $\prod_i(z_i - z_0)$. Similarly a quasiparticle with charge $-e/q$ can be formed by removing a flux line, which can be represented approximately by the operator $\prod_i(\partial/\partial z_i - qz_0/l_0^2)$ acting on the multinomial f.

Jain [92–94] has generalized this idea in the composite fermion model. Start with the wave function for p full Landau levels in some magnetic field, then attach $\pm 2m$ flux quanta to each electron, multiplying the wave function by $\prod(z_i - z_j)^{\pm 2m}$. Project the result back onto the lowest Landau level. This gives a filling factor, the number of electrons per flux quantum, equal to $p/(2mp \pm 1)$. For $m = 1$ this gives the two sequences tending to $1/2$ which are prominent in the experimental data, $1/3, 2/5, 3/7, \ldots$ and its particle-hole conjugate. For $m = 2$ the sequences of odd-denominator fractions tend to $1/4$. Further manipulation with particle-hole conjugation gives other fractions like the sequences tending to $3/4$.

The arguments that explain the integer quantum Hall effect are rather general, and have to be reconciled with the widespread occurrence of the fractional effect. After Laughlin's [90] explanation of the fractional quantum Hall effect was generally accepted, it was suggested that the ground state should have a discrete broken symmetry [74,95,96]. If there are q equivalent ground states, then, in Laughlin's argument for the integer effect [58], q flux quanta have to be introduced before the wave function returns to its original form. If only one flux quantum is introduced a different ground state is obtained.

This seemed to be contradicted by the apparent nondegeneracy of Laughlin's trial wave function for electrons confined to a disk [90], and by

the proved nondegeneracy of the wave function for electrons on a sphere [91]. However, one cannot do a measurement of the quantum Hall effect on the sphere without introducing probes of some sort through the surface, even if magnetic monopoles are available to provide a uniform magnetic field through the surface, and once probes are introduced one must worry about boundary effects. We have analysed this problem for the sphere [97] and the annulus [98]. For an annulus threaded by flux Φ the wave function takes the form

$$\left(\prod_{i=1}^{N} z_i^{n_1} |z_i|^{-\Phi e/h}\right) \left(\prod_{(ij)} (z_i - z_j)^q\right) \exp\left(-\sum_i |z_i|^2/4l_0^2\right), \qquad (4.19)$$

which is a superposition of determinants made up from one-electron wave functions of the form given in equation (4.8). The powers of $|z_i|$ in the prefactor of the exponential range from $n_1 - \Phi e/h$ to $n_1 - \Phi e/h + q(N-1)$, and so the electrons are concentrated in the range

$$2(n_1 - \Phi e/h)l_0^2 < |z|^2 < 2(n_1 - \Phi e/h + qN)l_0^2. \qquad (4.20)$$

The effect of increasing Φ adiabatically by h/e is to pull the whole electron distribution inwards, producing a state which is different from, and inaccessible from, the new quasiequilibrium state, which is obtained from equation (4.19) by simultaneously increasing $\Phi e/h$ and n_1 by unity. It is only when Φ has been increased by q that the new quasiequilibrium state can be reached by transferring one electron from the inner edge to the reservoir and one electron from the outer edge to the reservoir.

In this picture, whether applied to the annulus connected to two different reservoirs, or to the torus which was used to related Hall conductance to Chern numbers, the fractionally charged quasiparticles serve as point topological defects whose migration across the system can enable transitions to occur between these different ground states [98, 99]. An excitation with fractional charge e/q (quasihole) located at z_0 can be generated for the annulus by multiplying the wave function (4.19) by $A(z_0)\prod_i(z_i - z_0)$, where A is a normalization factor. When z_0 is very large, A is $1/(-z_0)^N$, and this factor does not change the wave function. For $z_0 = 0$ it increases n_1 by unity, pushing the electron distribution outwards, so that there is an extra charge e/q on the inner edge and $-e/q$ on the outer edge. This state can be reached by moving z_0 continuously from the outside of the annulus to the inside. While z_0 is in the interior of the annulus there is a quasihole of charge e/q crossing the annulus with some compensating charge density on each of the edges.

4.6   Skyrmions

In most of the discussion of the fractional quantum Hall effect up to this point it has been assumed that there is a single Landau level involved, with a single spin orientation. In certain systems this assumption is inappropriate, and two or more coupled Landau levels have to be considered. In most systems the Zeeman energy is relatively small, and in some systems it can be made particularly small, so the assumption that the spins of the electrons are completely polarized is inappropriate. A similar situation arises when the Zeeman energy is large, but a bilayer consisting of two adjacent layers of electrons is formed. In this case a pseudospin can be used to describe in which layer an electron lies.

Halperin [100] introduced a modification of Laughlin's wave function (4.18) which allows for equal occupation of the two spin or peseudospin states. This has the form

$$f_{\{m,m,n\}}(z_1, z_2, \ \ldots; z_{1'}, \ \ldots) =$$
$$\left[\prod_{i<j}^{N}(z_j - z_i)^m \right] \left[\prod_{i'<j'}^{N} (z_{j'} - z_{i'})^m \right] \left[\prod_{i,i'}^{N}(z_{i'} - z_i)^n \right], \qquad (4.21)$$

where the unprimed variables are the coordinates of spin up electrons and the the primed variables are the coordinates of spin down particles. Power counting shows that this gives the occupation number ν, the ratio of the total number of electrons to the number of flux quanta in the area they occupy, as $\nu = 2/(m + n)$. The exponents of $(z_j - z_i)$ and $(z_{j'} - z_{i'})$ must be the same, or else the spin up and spin down electrons will be occupying different areas, so this cannot be used to describe partial spin polarization.

States of the form $\{m, m, m\}$ in equation (4.21) are particularly interesting. For $m = 1$ this gives a singly occupied Landau level rotated by $\pi/2$ in pseudospin space, and for $m = 3, 5, \ \ldots$ it gives the Laughlin wave functions (4.18) rotated by $\pi/2$ in pseudospin space. All the pseudospins are aligned, so this is a pseudospin, or, for the case of vanishing Zeeman splitting, real spin ferromagnet. In regular Heisenberg ferromagnets, in two dimensions or in three, the lowest-lying part of the excitation spectrum consists of spin waves, in which a single spin is reversed. In two dimensions a skyrmion excitation is possible. This is a texture in the r, θ plane such as

$$S_x = \sin f(r) \cos \theta, \quad S_y = \sin f(r) \sin \theta, \quad S_z = \cos f(r), \qquad (4.22)$$

with $f(r)$ a function that is zero at infinity and goes smoothly to π at the origin. This texture gives a mapping of the plane (plus the point at infinity) onto the sphere of spin directions that wraps round the sphere and has a unit topological quantum number. The winding number is given by equation (1.3), with the integral over the unit sphere replaced by an

integral over the infinite plane. In the case given in equation (4.22) the winding number is $N_w = 1$. For the Heisenberg model the skyrmions have rather a high energy, but it has been argued that for the quantum Hall systems that are being considered in this section the skyrmion should give the lowest excitations for $\nu = 1$ [101–104]. As is shown in the work of Fertig *et al.* [105], the topological charge, the winding number, gives the electric charge of the excitation.

Several experimental groups have seen effects that seem to be due to such skyrmions at $\nu = 1$. Barrett *et al.* [106] and Aifer *et al.* [108] found a strong peak in the spin polarizability at $\nu = 1$ Schmeller *et al.* [107] found multiple spin flips (as many as 7) accompanying elementary excitations for $\nu = 1$, but not for $\nu = 1/3$ or $1/5$. Bayot *et al.* [109] found an enhancement of the specific heat, which are attributed to the skyrmion excitations at $\nu = 1$. Maude *et al.* [110] made electrical transport measurements, and found an enhanced spin gap even for $g = 0$, as would be expected if the excitations were skyrmionic.

5 Topological phase transitions

5.1 The vortex induced transition in superfluid helium films

In the conventional phase transitions there is usually some order parameter which has a nonzero value in the low temperature phase and is zero in the high temperature phase. The susceptibility associated with this order parameter diverges at a critical point, and this gives singularities in other quantities such as the specific heat. The two-dimensional Ising model was the first model solved in these terms, but renormalization group allowed good accounts of three-dimensional models such as the Ising, planar spin and Heisenberg models to be given. Also the theory of conformal invariance gave results for critical exponents in other discrete two-dimensional models such as the three-state Potts model. In such conventional continuous (second order) phase transitions, fluctuations away from the equilibrium value of the order parameter decay exponentially with a negative exponent whose value is the ratio of the distance to a temperature-dependent correlation length. At the critical point where the phase transition occurs the correlation length is infinite, and the correlations fall off as a power of the distance, rather than as an exponential function.

It was known from the work of Landau and Peierls [111] in the 1930s that two-dimensional systems with continuous order parameters could not have conventional long range order, and this was established rigorously by Mermin and Wagner [112] and by Hohenberg [113]. The possibility of some other type of order was not ruled out – it was known, for example, that a solid could have algebraic divergence of its X-ray peaks at the Bragg angles

[114], or it might have orientational long-range order without positional long-range order.

Onsager [1] and Feynman [20] suggested that the phase transition in bulk ^{4}He from the superfluid to the normal state might not be characterized by the disappearance of the magnitude $|\Psi_0|$ of the condensate wave function, but rather by a loss of phase coherence produced by a tangle of vortex lines. This does not fit well with modern renormalization group ideas, whereas the observed critical exponents of this transition do fit the renormalization group results for a two-component vector order parameter rather well. The results for a three-dimensional system with a two-component vector lie smoothly between the results for the Ising model and the Heisenberg model. Also it is technically difficult to handle a phase transition driven by vortex lines, because it involves a model like a directed polymer with a long-range interaction.

In two dimensions the situation is much simpler, since vortices are just point defects with a long-ranged interaction between them. We refer to a phase transition driven by such defects as a topological phase transition. Helium films provide the simplest example of a topological phase transition. The order parameter, the condensate wave function, is a single complex function of the two-dimensional position, so the system is equivalent to a planar spin model in two dimensions. The singularities of this system, the vortices, which play a vital role in this theory, are simply point singularities where the order parameter goes to zero, and around which the phase changes by $\pm 2\pi$. This type of system was studied in the work of Berezinskii [115,116] and of Kosterlitz and Thouless [117,118].

The development of the theory is simplified if we consider a lattice gas model of the superfluid, in which there is a magnitude $|\Psi_i|$ and a phase θ_i of the order parameter associated with each site i of the lattice. At low temperatures there will only be small fluctuations of the magnitude $|\Psi_i|$, and the phase fluctuations will be controlled by an energy term of the form

$$\frac{\mathcal{H}}{k_\mathrm{B}T} = -K \sum_{(ij)} \cos(\theta_i - \theta_j), \tag{5.1}$$

where (ij) denotes a pair of nearest neighbor sites. In the Gaussian approximation, the cosine in this formula is replaced by its quadratic approximation, to get the exponent of the Boltzmann factor as

$$-\frac{\mathcal{H}}{k_\mathrm{B}T} = -\frac{\mathcal{H}_0}{k_\mathrm{B}T} - \frac{1}{2}K \sum_{(ij)} (\theta_i - \theta_j)^2. \tag{5.2}$$

The angles θ_j are then written as their Fourier transforms, proportional to $\int \mathrm{d}^2k\, c(\mathbf{k}) \exp(i\mathbf{k}\cdot\mathbf{R}_j)$, so that equation (5.2) becomes a diagonal quadratic form in the Fourier components $c(\mathbf{k})$. The correlation function for $\cos\theta_i$ can

then be expressed in terms of an integral of an exponential whose exponent is quadratic and linear in these Fourier components, and the evaluation of this integral gives the correlation function as

$$\langle \cos\theta_i \cos\theta_j \rangle = \frac{1}{2}\exp\left[-\frac{a^2}{4\pi^2}\int\int \mathrm{d}^2 q \frac{1 - \cos(\mathbf{q}.\mathbf{R}_{ij})}{4 - 2\cos(q_x a) - 2\cos(q_y a)} \right], \quad (5.3)$$

where I have taken the lattice to be a square lattice of side a, and $\mathbf{R}_{ij}$ is the vector between the sites i and j. For large values of R_{ij} the integrand is of order $1/q^2 a^2$ for q between $1/R_{ij}$ and $1/a$, so the integral depends logarithmically on the ratio of these two quantities, and the result

$$\langle \cos\theta_i \cos\theta_j \rangle \approx \left(\frac{a}{R_{ij}} \right)^{\eta(T)}, \quad (5.4)$$

with

$$\eta(T) = \frac{1}{2\pi K}, \quad (5.5)$$

is obtained. This power law behavior of the correlation function at all low temperatures was discussed by Wegner [119] and by Jancovici [120], and means that the whole low-temperature region can be regarded as a critical line. For the superfluid the energy associated with variation of the phase can be written as

$$\mathcal{H} = \frac{1}{2}\int\int \mathrm{d}^2 r \rho_s \frac{\hbar^2}{m^2}(\mathbf{grad}\theta)^2, \quad (5.6)$$

and comparison of equations (5.1, 5.5) and (5.6) shows that the correlation exponent is given by

$$\eta(T) = \frac{m^2 k_B T}{2\pi\hbar^2 \rho_s}, \quad (5.7)$$

where ρ_s is the superfluid density per unit area. This algebraic fall-off of the correlation function implies that the usual order parameter, related to the infinite distance limit of the correlation function, must be zero.

In accordance with the discussion of superfluid flow in Section 2.1, the unbounded fluctuation of the relative phase at large distances implied by equation (5.4) will not be enough to destroy superfluid flow in an annulus, which is represented by a texture of the phase in which the phase changes by a multiple of 2π round the annulus. The phase can only lose this twist by the passage of vortices across the system, and it turns out that, in the thermodynamic limit, there is an infinite barrier to the creation of the necessary vortices. The energy of an isolated vortex in a system of linear dimension L is

$$E_v = \frac{1}{2}\rho_s \int_{a_0}^{L} 2\pi r\,\mathrm{d}r \frac{\hbar^2}{m^2}\frac{1}{r^2} = \frac{\pi\rho_s\hbar^2}{m^2}\ln\frac{L}{a_0}, \quad (5.8)$$

where a_0 is the vortex core radius. There is also an entropy associated with the possible positions of a vortex. If the number of possible positions in a system of area L^2 is $L^2/\pi a_1^2$, then the entropy per vortex is

$$S_{\mathrm{v}} = k_{\mathrm{B}} \ln(L^2/a_1^2) . \qquad (5.9)$$

At low temperatures the free energy per vortex is positive, and depends logarithmically on the size of the system, so the concentration of free vortices is zero in the thermodynamic limit at low temperatures. Above the temperature given by $k_{\mathrm{B}}T = \pi\rho_s\hbar^2/2m^2$ the logarithmic term in the free energy of a vortex changes sign, so free vortices appear, and supercurrents are unstable. This argument leads to the conclusion obtained by Nelson and Kosterlitz [121] that a necessary condition for the stability of superfluid flow is

$$\rho_{\mathrm{s}} \geq \frac{2m^2 k_{\mathrm{B}}T}{\pi\hbar^2} , \qquad (5.10)$$

or, from equation (5.7),

$$\eta(T) \leq 1/4 . \qquad (5.11)$$

In the more detailed theory this inequality becomes an equality at the transition temperature, and the critical line ends at $\eta = 1/4$.

We described this type of phase transition as a *topological phase transition* because the most prominent feature of the ordered low-temperature phase is the absence of topological defects and the stability of states with nonzero topological quantum number in the thermodynamic limit, rather than the existence of a conventional order parameter.

In the detailed theory [118], the effect of creation of bound pairs of positive and negative vortices on the superfluid density was taken into account. The energy of interaction of two vortices with quantum numbers n_1, n_2, giving phases $\theta_1(\mathbf{r}), \theta_2(\mathbf{r})$, is

$$\int \int \mathrm{d}^2 r \frac{\rho_s\hbar^2}{m^2} \mathbf{grad}\theta_1(\mathbf{r}) \cdot \mathbf{grad}\theta_1(\mathbf{r}) = -\frac{2\pi\rho_s\hbar^2}{m^2} n_1 n_2 \int_{a_0}^{R_{12}} \frac{\mathrm{d}r}{r}$$

$$= -\frac{2\pi\rho_s\hbar^2}{m^2} n_1 n_2 \ln \frac{R_{12}}{a_0} \cdot \quad (5.12)$$

The vortices therefore behave like a Coulomb gas (with a two-dimensional Coulomb interaction) of classical charges. The bound vortex pairs are like classical molecules. The Hamiltonian for this classical Coulomb gas can be written in lattice gas form as

$$\frac{H}{k_{\mathrm{B}}T} = -\pi K \sum_{i \neq j} n_i n_j \ln \frac{\mathbf{r}_i - \mathbf{r}_j}{\tau} + \frac{\mu}{k_{\mathrm{B}}T} \sum_i n_i^2 + \pi K \sum_{i,j} n_i n_j \ln \frac{L}{\tau} \cdot \quad (5.13)$$

Here K is $\rho_s\hbar^2/m^2 k_{\mathrm{B}}T$, τ is the lattice spacing, of order a_0, μ is a chemical potential to represent the short-range contributions to the vortex energy,

and the final term represents the long range contributions to the energy in a system of linear dimension L when $\sum n_i$ is nonzero. The effective dielectric constant K is assumed to be a function of distance, because the smaller vortex pairs are polarized by the fields acting between more distant pairs. A crude treatment of the polarizability gives it as

$$\langle R_{12}^2 \rangle = \frac{\int_{r_0}^{\infty} R^3 \mathrm{d}R \, \exp[-2\pi K \ln(R_{12}/\tau)]}{\int_{r_0}^{\infty} R \mathrm{d}R \, \exp[-2\pi K \ln(R_{12}/\tau)]} = \tau^2 \frac{\pi K - 1}{\pi K - 2} \, . \tag{5.14}$$

This diverges at $K = 2/\pi$, where $\eta(T) = 1/4$, which is the bound given in the expression (5.11); with divergent polarizability there is no barrier to the dissociation of vortex pairs.

Renormalization arguments originally developed by Anderson and Yuval [122] were applied by Kosterlitz [123] to this problem of interacting vortices or interacting charges in two dimensions. This method was developed further by José et al. [124]. The basic idea of this method is to renormalize the free vortex (or free charge) fugacity $y(L) = \exp(-\mu/k_\mathrm{B}T)$ at each length scale L in terms of the dimensionless superfluid density (or reciprocal of the dielectric constant) $K(L)$. The dielectric constant on the length scale L is determined by those dipoles whose size is less than L, so the change of the dielectric constant when the length scale is varied is proportional to the square of the concentration of free charges found on length scales L. This leads to the first of the Kosterlitz equations

$$L\frac{\mathrm{d}K^{-1}(L)}{\mathrm{d}L} = 4\pi^3 y^2(L) + O(y^4) \, . \tag{5.15}$$

The self-energy of a charge (or vortex) on a length scale L depends on the polarizability of the medium on that length scale and on the concentration of free vortices, so that leads to the second Kosterlitz equation

$$L\frac{\mathrm{d}y(L)}{\mathrm{d}L} = [2 - \pi K(L)]y(L) + O(y^3) \, . \tag{5.16}$$

More detailed derivations of these equations can be found in the review by Nelson [125].

If the concentration y of defects (charges or vortices) is sufficiently small that higher order terms in y can be ignored, the two equations can be combined to get

$$\frac{\mathrm{d}K(L)}{\mathrm{d}y(L)} = -\frac{4\pi^3 y K^2}{2 - \pi K} \, , \tag{5.17}$$

which can immediately be integrated to get the relations between $K(L)$ and $y(L)$

$$\ln\frac{\pi K(L)}{2} + \frac{2}{\pi K(L)} = 2\pi^2 y^2(L) + 1 + \alpha \, , \tag{5.18}$$
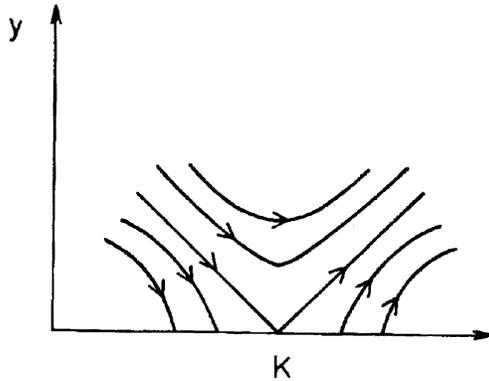
Fig. 4. Flow diagram for the renormalization of K and y from equations (5.15) and (5.16).

where α is a constant of integration. These solutions, which represent the renormalization flow of K and y, are shown in Figure 4. Those curves for which α is positive flow into solutions for large L with y tending to zero, whereas if α is negative the flow may initially be to smaller concentrations y, but eventually goes to larger concentrations, so that there are free defects on a large length scale. The explicit equation (5.18) for y as a function of K can be substituted back into equation (5.15) to get a set of expressions for K as a function of L.

To use this scaling relation to obtain long-range properties it is necessary to start with known properties at length scales of order a_0, the vortex core radius. The superfluid density $K(a_0)$ and the vortex fugacity $y(a_0)$, unmodified by the specific two-dimensional properties we have been discussing, will both be functions of temperature, as will the core radius itself. From where this trajectory $K(a_0(T), T)$, $y(a_0(T), T)$ cuts the flow lines (5.18), the long-range superfluid density $K(\infty, T)$ for $T < T_c$ and the transition temperature T_c can be read off. The most striking result is the one already mentioned [121], that, at the temperature where free vortices first appear and the superfluidity is destroyed, $K(\infty, T_c)$ is always $2/\pi$, and the superfluid density always has the value

$$\rho_s(T_c) = \frac{2m^2 k_B T_c}{\pi \hbar^2} \, . \tag{5.19}$$

This relation was verified by Rudnick [126] studying the propagation of third sound, the waves which propagate on superfluid films, and by Bishop and Reppy [127, 128] using an Andronikashvilii oscillator constructed of a roll of mylar film with a thin coating of liquid helium. Both these experiments

use a nonzero measuring frequency; this also implies a finite length scale of the order of the wavelength of third sound at that frequency. The theory used to interpret such measurements at nonzero frequency measurements was developed by Amegaokar *et al.* [129]. At such nonzero frequencies the transition is rounded and pushed up to higher temperatures, and a good fit to the experimental data was obtained with this theory, and an extrapolation can be made to the discontinuity of superfluid density and the transition temperature at zero frequency or infinite length scale. Later work by McQueeney *et al.* [130] on solutions of ^{3}He in ^{4}He carried this relation down to rather low temperatures.

5.2   Two-dimensional magnetic systems

Theoretical studies of critical phenomena usually relate the universality class to an equivalent magnetic model. The liquid-vapor critical point and the magnetic system with a single preferred axis of magnetization both are in the same universality class as the Ising model. The superfluid transition, with an order parameter that is a complex scalar quantity, resembles a magnetic system with a preferred plane of magnetization but no preferred direction in the plane. This is the planar spin model.

The three-dimensional planar spin model is similar in many ways to other three-dimensional magnetic models, intermediate in its critical exponents between the Ising model with a single preferred axis and the Heisenberg model with no preferred direction of magnetization. In two dimensions, however, the planar spin model has very special properties. There is no net magnetization in zero field, and the spin-spin correlation function tends to zero at large distances except at zero temperature. However, below the critical temperature for vortex pair unbinding the spin-spin correlation function decays algebraically with distance, rather than exponentially, and the spin-wave stiffness is nonzero. For spins on a lattice one can define vortices uniquely, as points on the dual lattice around which the adjacent spins rotate through an angle of $\pm 2\pi$, if one defines the angle between two neighboring spins to be in the range from $-\pi$ to π. Isolated vortices have an energy that depends logarithmically on the size of the system, and vortex-antivortex pairs have an energy that depends logarithmically on their separation. The phase transition is driven by the dissociation of vortex-antivortex pairs, and is precisely analogous to the superfluid transition discussed in the previous subsection. The spin-wave stiffness is the quantity that is analogous to the superfluid density.

When the rotational symmetry in the plane is broken various thing can happen. If there is a two-fold symmetry – two equivalent minima in the energy at an angle π apart – the magnetic system has the critical behavior of the Ising model. With a three-fold symmetry axis, so that the minima in energy are at an angle $2\pi/3$ apart, the critical behavior is that of the three-state Potts model. If the symmetry axis perpendicular to the plane has five-fold or higher symmetry, as it will if the anisotropy is due to crystal fields in a triangular lattice, then the anisotropy can be shown to lock the magnetization into one of the preferred directions at sufficiently low temperatures, but at intermediate temperatures the anisotropy becomes irrelevant, so that the system behaves like an isotropic planar spin system, with algebraically decaying long-range order, whose exponent changes continuously with temperature until it reaches the critical value at which free vortices can form spontaneously [124]. The case of anisotropy with four equivalent axes in the plane lies on the boundary between these two types of behavior, and algebraically decaying order occurs only at one temperature, but the exponent at this temperature has a nonuniversal value.

The topological nature of the planar spin transition is shown also by considering what happens in the Heisenberg model in two dimensions, with the spin free to point in any direction in space [118, 131]. In this case there is no longer a metastable vortex, a singular point around which the spin rotates by an angle 2π, because the spin is free to be directed anywhere on the sphere. This allows the vortex to be replaced by a texture in which the spins near the center of the texture are tilted out of the xy plane towards the z direction, so that the direction of the spins is a smooth function of position. The energy of this texture, instead of being proportional to $\ln(L/a_0)$, where a_0 is of the order of the lattice spacing, is proportional to $\ln(L/\xi_0)$, where ξ_0 is the length scale of this central region in which the spins are tilted out of the plane. Since ξ_0 is continuously variable this energy is continuously variable down to zero, and so there is no energy barrier to the creation or annihilation of vortex pairs. There is no phase transition for the two-dimensional Heisenberg model at nonzero temperature.

5.3   Topological order in solids

A solid can be characterized in a number of different ways. It has positional long-range order which is characterized by the existence of a reciprocal lattice and shown up by the sharp Bragg peaks in X-ray scattering. It has orientational long range order, in that the directions from one atom to its nearest neighbors in one part of the crystal are correlated with those in some other part of the crystal. It is rigid, in the sense that it resists a shearing stress and has infinite viscosity – it may yield to a shearing stress, but the rate of yield in an ideal solid rises much slower than linearly with the applied stress.

Associated with the symmetry breaking from continuous to discrete translational symmetry there is an order parameter which is the local displacement of the actual lattice from the ideal lattice, which is the quantity whose Fourier transform determines the sharpness of the Bragg peaks. In a two-dimensional system the peaks become algebraic rather than δ-function peaks [114], because thermal fluctuations of this order parameter grow logarithmically with distance.

The topological defects associated with this order parameter are *dislocations*. A typical edge dislocation in a square lattice is shown in Figure 5. In three dimensions this can be formed by the Volterra process, in which a half-plane of atoms bounded by the dislocation core is removed, and then the atoms are rejoined across the cut. If we define the vector **u** to be the displacement of the local lattice relative to the closest point of the ideal lattice, then the gradient of **u** round a closed loop gives a lattice vector, which is the Burgers vector of the dislocation. In the case of an edge dislocation this vector is in the plane perpendicular to the core. One can find the Burgers vector by mapping a closed path round the dislocation core onto an ideal lattice; the Burgers vector is the amount by which the map of the path fails to close. The dislocation is known as a screw dislocation if the Burgers vector is parallel to the dilocation core, but this is not relevant to the two-dimensional case, where the core is a point defect, and the Burgers vector must lie in the plane. The topological description of dislocations has been given in a paper by Kléman [132].
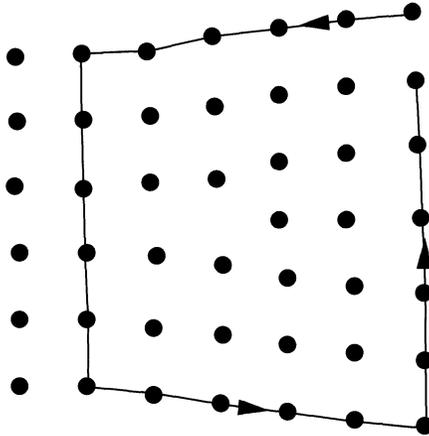
Fig. 5. Edge dislocation formed by the Volterra process. A Burgers circuit is shown, and the Burgers vector is the amount by which it fails to close.

The rigidity of a solid is unrelated to the sharpness of the Bragg peaks or the thermal fluctuations of the order parameter, but depends on the absence of dislocations in the crystal lattice [133]. Dislocations can move under the influence of shearing stress to relieve the stress, and if this process is thermally activated under conditions of low shearing stress it will lead to a viscous flow with a rate of strain proportional to stress. The dislocation produces a strain field, both longitudinal and transverse, which falls off linearly with the reciprocal of the distance from the core. As a result, the energy of a single dislocation depends logarithmically on the size of the system in two dimensions, and the interaction between two dislocations also depends logarithmically on the distance between them. However, this interaction also depends on the Burgers vectors of the dislocations; the logarithmic interaction is proportional to the scalar product of the two Burgers vectors, but there is an additional term independent of distance that is proportional to the projection of the Burgers vectors in the direction of the displacement of one dislocation relative to the other.

The theory of dislocation mediated melting in two dimensions was worked out by Young [134] and by Halperin and Nelson [135]. The theory is superficially similar to the theory of superfluidity in two dimensions, but there are important differences introduced by the vector character of the Burgers vector that characterizes the dislocations. For example, in place of equation (5.10) there is the more complicated relation

$$\frac{\mu(\mu + \lambda)}{2\mu + \lambda} > \frac{4\pi k_B T}{a_0^2} \,, \qquad (5.20)$$

where μ, λ are the elastic moduli characterizing transverse and isotropic strain and a_0 is the lattice constant. This gives a lower bound on the rigidity modulus μ in the solid phase, just as equation (5.10) gives a lower bound for the superfluid density in the superfluid phase. The same combination of elastic moduli is involved in the generalization of the Kosterlitz scaling relations (5.15) and (5.16) [134,136]. Instead of a single exponent η characterizing the algebraic decay of correlations there is a set of exponents, one for each Bragg peak.

Above the dislocation unbinding temperature there may still be orientational order [136], but it now has an algebraic decay with distance, instead of having a nonzero long-range limit as it has below the transition temperature. Orientational order can also be defined in terms of a topological property. At each point in the dislocated crystal, except very close to the dislocation cores, the short range order of the atoms definies a local set of crystal axes. A parallel translation of these local axes can be made along a long closed path, and this defines a continuous mapping of the path on the space of axis orientations. If this mapping cannot be shrunk to a point there must be a singularity enclosed by the path. Such singularities are

disclinations, similar to the disclinations found in liquid crystals and in the
A phase of superfluid ^{3}He. They can be constructed from the ideal lattice by
removing or adding a slice of atoms, for example a $\pi/3$ sector in a triangular
lattice, or a $\pi/2$ sector in a square lattice, and rejoining the edges smoothly.
At the center of a disclination in a triangular lattice a lattice site has five or
seven nearest neighbors instead of the six which is usual. In a normal solid
the elastic energy of a disclination is very high, proportional to the area
of the system, but once the dislocation unbinding transition has occurred
this energy depends logarithmically on the size of the system, so that the
statistical mechanics of disclinations in this region is very similar to the sta-
tistical mechanics of vortices in a superfluid. Nelson and Halperin [135,136]
called this phase *hexatic.* Dislocations can be regarded as tightly bound
pairs of disclinations of opposite sign, and the transition to an unoriented
liquid goes by the process of unbinding of disclination pairs. In the hexatic
phase the orientational order falls off algebraically with an exponent η_6, and
the transition occurs when η_6 has the critical value $1/4$, in agreement with
equation (5.11).

The most commonly studied form of two-dimensional solid is a ph-
ysisorbed layer, such as a noble gas adsorbed on a graphite or other solid
surface. In this case the substrate provides orientational ordering at all
temperatures. The adsorbate may be in registry with the substrate, but
it may also form its own crystal lattice incommensurate with that of the
substrate. In this case it can undergo a dislocation unbinding transition to
a fluid state, and this has been studied by Young [134].

Most systems studied by experiment or in computer simulations actually
undergo a first order phase transition to a fluid state before they reach the
bound given by equation (5.20). A possible exception is the two-dimensional
Coulomb solid, which was first found in experiments by Grimes and Adams
[137]. At least both experimental [138] and theoretical [139] studies have
shown a rigidity which comes very close to the theoretical bound given by
equation (5.20). A review both of the theory and of its possible applications
has been given by Strandburg [140].

5.4   Superconducting films and layered materials

The vortices in a charged superfluid generate a magnetic field, so that, as
we have seen in Section 2.1, it is not the circulation which is quantized
but the fluxoid, a combination of the circulation round a loop and the
magnetic flux enclosed by it. On length scales small compared with the
London penetration depth the vortices in a superconducting thin film are
very similar to those in a neutral superfluid thin film, but on length scales
larger than the penetration depth they are quantized flux lines. One result of
this is that the logarithmic interaction between a vortex and an antivortex is
cut off at a distance of the order of the penetration depth, to be replaced by

an interaction that falls off as the inverse cube of the distance, characteristic of the interaction between localized magnetic dipoles.

Since the logarithmic interaction cuts off at the penetration depth we supposed that vortex pair unbinding would occur at arbitrarily low temperatures, and that there should be no phase transition in superconducting thin films [118]. In fact, however, the London penetration depth in very thin films is inversely proportional to the film thickness, and can be rather large, comparable with the size of the sample. Therefore, as was argued by Beasley *et al.* [141] and by Doniach and Huberman [142], there can be a vortex unbinding transition in superconducting thin films. It will not be a sharp transition, but will have a rounding determined by the penetration depth or by the sample size, whichever is smaller. The resistivity and surface impedance of such superconducting films have properties which are characteristic of the vortex-unbinding transition, such as a current-voltage relation with an exponent that passes through the value $1/3$ at the transition, and this behavior has been observed in a number of experiments [143–145]. There is a review of the theory and its applications to superconductivity by Minnhagen [146].

For a superconducting film in a magnetic field a number of issues arise, which were first discussed by Huberman and Doniach [147] and by Fisher [148]. In an ideal superconductor the flux lines will form a lattice at low temperatures, but this lattice should either melt when it becomes unstable to the dissociation of dislocation pairs, or should have a first order phase transition at a lower temperature. Once the lattice is melted it may be in a hexatic phase, which would then become isotropic at higher temperatures, or it may go directly into an isotropic fluid phase if there is a first order transition.

The vortex-driven transition from the superconducting state has also been extensively studied in two-dimensional arrays of Josephson junctions. One way of fabricating such a device is to deposit a regular array of superconducting islands on a normal metal substrate, such as copper. At low temperatures, in the absence of an external magnetic field, not only will all the islands be superconducting, but their phases will be locked together by Josephson tunneling through the normal metal substrate. If an external magnetic field is applied normal to the plane of the array there will be diamagnetic currents in the superconducting array which will cause vortices to form in the junction array. Energetically it is far more favorable for a vortex to form in a normal region between neighboring superconducting islands, a point in the lattice dual to the lattice of islands, than in the superconducting islands themselves. Only the small Josephson currents between the islands are affected by the vortex, whereas a vortex in one of the superconducting islands would lead to suppression of the superconductivity in that island. Even in the absence of an external magnetic field vortex-antivortex pairs

can form on neighboring or close sites of the dual lattice. The phase of the superconducting order parameter provides an order parameter similar to that of the planar spin model, so the energy of a vortex pair depends logarithmically on the separation of the vortex from the antivortex. In principle there is a London penetration distance for the network, beyond which the magnetic field is screened, and the interaction between vortices is power law rather than logarithmic, but in practice the Josephson currents are too small for this to be important. The thermally driven dissociation of the vortex pairs therefore leads to a zero-field phase transition with the same properties as the other transitions discussed in this chapter, such as the superfluid transition in ^{4}He films and the melting of the hexatic phase in two dimensions.

At temperatures below this vortex-unbinding transition, current can be passed from one side of the network to the other with no resistance just as it can in a bulk superconductor. Above the unbinding transition there is a voltage generated when a current is passed. This voltage is produced by the free vortices moving to one edge of the array, across the current, and the antivortices moving to the other edge. These vortices that move to one edge of or the other are replaced by thermally excited vortex-antivortex pairs. Although the network cannot pass a supercurrent free of dissipation above this transition temperature, the islands of which it is composed still have the usual properties of small superconducting regions. Thus the superconducting order parameter still exists locally, although its phase is not locked across the sample. This is the same picture which is used to describe the vortex-driven transition in superconducting films and in superfluid films. This transition in superconducting networks was observed by Abraham *et al.* [149], who reported that when they studied the voltage-current relation they found $V \propto I^\alpha$, where the exponent α is greater than 3 in the superconducting phase, and drops sharply to unity at the transition temperature. This is the same behavior that was observed in superconducting films.

6   The A phase of superfluid ^{3}He

6.1   Vortices in the A phase

The atoms of ^{3}He are fermions, with nuclear spin 1/2, and the superfluid condensate is formed from a P-state, which implies that the nuclear spins must also be coupled together in a triplet state. This P-state pairing results in a very rich behavior of the order parameter, with far more possibilities than exist in superfluid ^{4}He or in conventional superconductors. The triplet pairing of the nuclear spins also allows nuclear magnetic resonance to be used as a tool to explore details of this behavior.

The A phase of the superfluid is stable in a region that extends along the solidification pressure (33 atm.) from 2.7 mK to 2.1 mK, and down to a point at 20 atm. and 2.5 mK on the phase boundary between normal fluid and superfluid. In this A phase the orbital state of the pair is described by a state with $M_L = 1$ along some axis which is denoted by $\mathbf{l}$. The spin state of the pair has $M_S = 0$ along some other axis which is denoted by $\mathbf{d}$. The hyperfine magnetic coupling between the nuclear spins tends to align the vectors $\mathbf{l}$ and $\mathbf{d}$ parallel or antiparallel to one another, but this hyperfine coupling is relatively weak, and under many circumstances they can take up directions which are independent of one another. This superfluid phase is therefore anisotropic in space, with two distinguished axes, or one axis if the directions of $\mathbf{l}$ and $\mathbf{d}$ coincide.

Early treatments of the topological properties of the A phase of ^{3}He, together with discussions of liquid crystals, were given in the papers of Toulouse and Kléman [150] and of Volovik and Mineev [151]. There is a detailed treatment in the book by Volovik [152].

To understand how this type of order parameter affects the topological properties of the phase we will first of all neglect the spin part of the order parameter. The axis of quantization of the orbital angular momentum of the pair can be chosen in any direction in space. Since this is an $M_L = 1$ state, rotation about this axis changes the phase, and we can represent phase and direction together by taking an element of the group $SO(3)$; in concrete terms by choosing three Euler angles to set direction and phase. The group $SO(3)$ is not simply connected, since a path which gives a rotation by 2π cannot be shrunk to zero. It is more convenient therefore to work with the simply connected group $SU(2)$ which gives a twofold covering of $SO(3)$. There are two convenient ways of representing the topology of $SU(2)$. One is in terms of the surface of a 3-sphere – a unit vector in the four-dimensional space of real coefficients of the unit matrix and the three Pauli matrices. Another way is in terms of a point in the interior of a sphere of radius 2π. The distance of the point from the center of the sphere gives the angle of rotation about an axis whose direction is given by the vector from the center to the point; in this representation all points on the surface of the sphere are equivalent to a 2π rotation.

Each possible orientation of the $\mathbf{l}$ axis and phase is represented by two points in $SU(2)$, in accordance with the double covering of $SO(3)$. One is separated from the other by a rotation of 2π, so they are represented by points lying on the same diameter of the 2-sphere separated by a distance 2π. Closed paths in this space can either be trivial loops in the space, or paths leading from one point to its equivalent point. The homotopy group π_1 is therefore equivalent to the group Z_2, the group given by multiplication of ± 1, or addition of the integers 0 and 1 modulo 2.
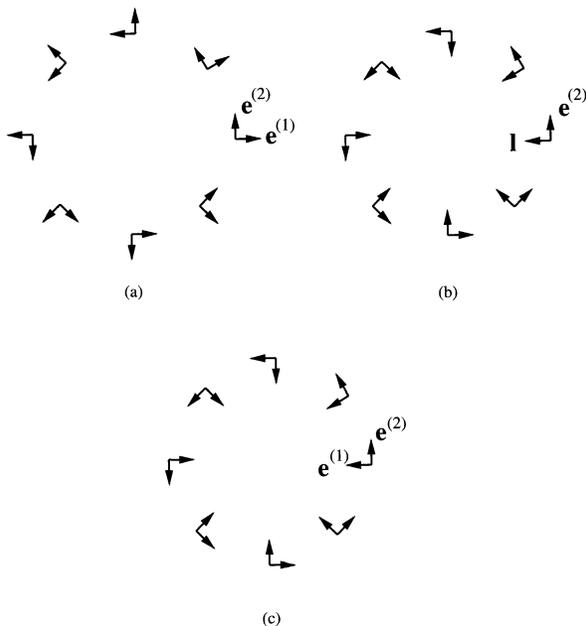
Fig. 6. This shows the continuous unwinding of a π twist in the order parameter for the A phase of ^{3}He by rotation about the tangential direction. In (a) and (c) the directions of twist are reversed, and in the intermediate case (b) it is obvious that there is no superfluid circulation.

It is clear that such a quantum number, with values 0 and 1, cannot correspond to a classical variable like circulation. For the A phase the superfluid velocity is given by

$$\mathbf{v_s} = (\hbar/2m_3) \sum_{j=1}^{3} e_j^{(1)} \mathbf{grad}\, e_j^{(2)} , \tag{6.1}$$

where $\mathbf{e}^{(1)}, \mathbf{e}^{(2)}$ are the two coordinate axes perpendicular to $\mathbf{l}$. When $\mathbf{l}$ has constant direction this corresponds to equation (2.1), with m_4 is replaced by $2m_3$. Anderson and Toulouse [153] argued that circulation can always be changed in a continuous way. An example of how this can be done is shon in Figure 6. Figure 6a shows the variation of the orientation of the axis-system when there is unit circulation round a channel. The direction of $\mathbf{l}$ is shown as out of the plane. If the orientation of the axis-system is slowly rotated, for example about the tangential direction, the circulation is reduced and can be reversed. Figure 6b shows the situation when a $\pi/2$ rotation has been made, so that $\mathbf{l}$ is along the inward radius and $\mathbf{e}^{(1)}$ is in
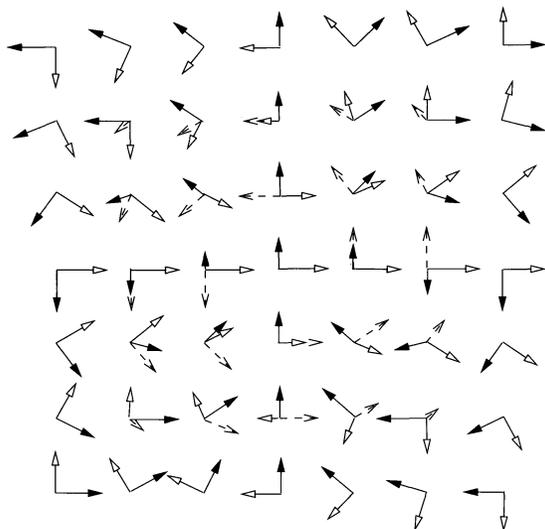
Fig. 7. Possible orientation of the order parameter for a nonsingular doubly quantized vortex in the A phase of ^{3}He. The directions of $\mathbf{e}^{(1)}$ and $\mathbf{e}^{(2)}$ are shown with solid and hollow arrowheads, and of $\mathbf{l}$ with a broken line.

the axial direction. In this case equation (6.1) shows that the flow velocity is zero. With the further rotation by $\pi/2$ shown in Figure 6c the circulation has been reversed. This process shows why the topological quantum number -1 is not distinct from 1 for the A phase.

Two important consequences follow from this. The first is that circulation is not quantized around channels unless some constraint is placed on the direction of $\mathbf{l}$ by the boundaries, as it is, for example in a thin film where $\mathbf{l}$ is constrained to be perpendicular to the film. The second is that when vortices do occur, for example in a rotating system, they can form an array of nonsingular doubly quantized vortices, where the circulation is $2h/2m_3$ around a path far from the vortex, where the path in the angle-axis space is a diameter of the 2-sphere or an equator of the 3-sphere. Closer in to the axis the path followed by the axis-system can shrink, so there is no singularity at the axis, where this path in order parameter space has shrunk to a point. An isolated doubly quantized vortex could have the three axes represented by the columns of the matrix

$$\begin{pmatrix} \cos^2\phi + \sin^2\phi\cos f & \sin\phi\cos\phi(1-\cos f) & \sin\phi\sin f \\ \sin\phi\cos\phi(1-\cos f) & \cos^2\phi\cos f + \sin^2\phi & -\cos\phi\sin f \\ -\sin\phi\sin f & \cos\phi\sin f & \cos f \end{pmatrix}, \quad (6.2)$$

where ϕ is the azimuthal angle measured from the vortex center, and $f(r)$ is a function of the distance r from the vortex center that goes smoothly from 0 at $r = 0$ to π for $r > a_0$. In such a configuration the axis system rotates by an angle 2π when one goes between two points on opposite sides of the vortex, but the rotation is about the z-axis for a path that avoids the vortex center, while it is about an axis in the xy plane for a path that goes straight through the center. Such a nonsingluar double vortex is illustrated in Figure 7. These textures are characterized by topological quantum numbers, but the quantum numbers are stabilized not by their topological structure, but by the external symmetry-breaking perturbation such as the imposed rotation. Recent work by Parts *et al.* [154] has shown regions of singly and doubly quantized vortices coexisting in the same rotating containers of A-phase superfluid.

The nuclear spin of the pair adds an extra complication to the classification of line defects in the A phase. If the direction of quantization of the spin is rotated by an angle π about a perpendicular axis, the spin wave function changes sign. This sign change can compensate for the sign change which occurs for a $\pm\pi$ rotation about the $\mathbf{l}$ direction. There are therefore two more equivalent points for each pair of equivalent points in $SU(2)$. Since the combination of two $\pm\pi$ rotations of the orbital system about $\mathbf{l}$ and two π rotations of the real vector $\mathbf{d}$ gives a zero or 2π rotation of the orbital angle-axis system and zero rotation of the nuclear system, the homotopy group is isomorphic to the group Z_4, the group of multiplications of fourth roots of unity. The topological quantum numbers of linear defects can take on the four values, which are assigned the numbers $0, \pm 1/2, 1$. This assignment of the quantum numbers ensures that the familiar single vortex, in which the phase of the order parameter changes by 2π round the vortex, has quantum number unity.

The half-integer quantum numbers can correspond to vortices with circulation $\pm h/4m_3$, combined with disclinations in the direction of $\mathbf{d}$ by an angle π. Such a half-integer vortex is shown in Figure 8. As it is shown there, $\mathbf{l}$ is out of the plane, and the phase rotates by π on a path round the vortex. At the same time the direction $\mathbf{d}$ of spin quantization also rotates by π, so that the apparent discontinuity across the half-plane is removed by two factors of -1. The hyperfine energy of such a defect will be large, because $\mathbf{d}$ cannot match the direction of $\mathbf{l}$ at large distances.

A survey of the properties of vortices in superfluid ^{3}He can be found in a paper by Krusius [155].

6.2   Other defects and textures

In accordance with the discussion in Section 1.3, topologically stable interface defects exist only when the states on the two sides of the defect have a difference in discrete symmetry. At the $A - B$ phase boundary the two
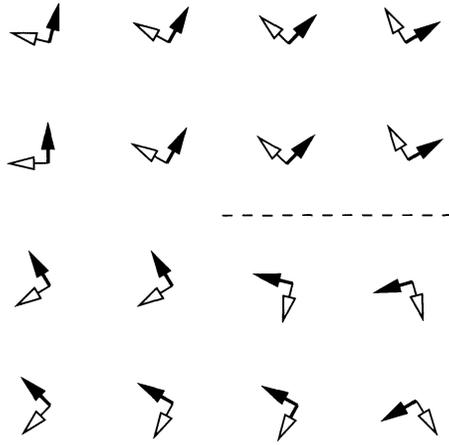
Fig. 8. Half-integer vortex for the A phase, with a π rotation of the orbital angular momentum or phase variable combined with a π disclination of the spin axis.

phases have the same free energy, but different and incompatible discrete symmetries, such that one symmetry cannot pass smoothly into the other. The interface between them is therefore topologically stable.

One example of such a phase coexistence is provided by rotating B phase ^{3}He. It may be energetically favorable for the vortex cores to take advantage of the breaking of time reversal symmetry produced by the rotation to undergo a transition to the A phase. There have been both theoretical [156,157] and experimental [158] studies of this.

In general a vortex sheet, separating two regions of the same superfluid with different uniform velocities, is both topologically and energetically unstable, and will break up into a row of vortex lines, separated by smoothly joined regions between them. As was mentioned in Section 1.3, there is an exception to this in the A phase of ^{3}He, which has been the subject of theoretical and experimental work recently [10,11]. The vortex sheet that this work describes is stabilized by the locking of the orbital vector $\mathbf{l}$ to the spin orientation $\mathbf{d}$ parallel and antiparallel on the two sides of the vortex sheet, so that $\mathbf{l} \cdot \mathbf{d}$ has the values ± 1 on the two sides of the vortex sheet, and both of these orientations minimize the dipolar hyperfine energy. This is topologically stable, because the two different orientations of the orbital angular momentum relative to the spin give two disjoint stable regions for the order parameter. A texture that joins two regions of opposite orbital angular momentum, each with different flow velocities $v_1 = \hbar k_1/2m_3$, $v_2 = \hbar k_2/2m_3$, across a planar vortex sheet, could have the unit vectors $\mathbf{e}^{(1)}$, $\mathbf{e}^{(2)}$ given by

the column vectors

$$
\begin{pmatrix} (1-p)\cos(k_1 x) + p\cos(k_2 x) \\ -(1-p)\sin(k_1 x) + p\sin(k_2 x) \\ -2\sqrt{p(1-p)}\sin[\frac{1}{2}(k_2+k_1)x] \end{pmatrix} \ , \quad
\begin{pmatrix} (1-p)\sin(k_1 x) + p\sin(k_2 x) \\ (1-p)\cos(k_1 x) - p\cos(k_2 x) \\ 2\sqrt{p(1-p)}\cos[\frac{1}{2}(k_2+k_1)x] \end{pmatrix} \ .
$$

$$(6.3)$$

Here p is a parameter that interpolates smoothly from zero to unity on the two sides of the vortex sheet, and the spin axis $\mathbf{d}$ is the same on the two sides of the sheet. The orbital anglular momentum vector $\mathbf{l}$ is the vector product of these two vectors, and reverses its direction as p goes from zero to unity. In the rotating systems that have been studied experimentally there seems to be a single vortex sheet that spirals through the system in such a way as to make the circulation density almost uniform, just as a vortex array gives an almost uniform circulation density in the more usual situation.

Point defects are characterized by the way the order parameter behaves on a closed surface, with the topology of a sphere, surrounding the defect. The possible topologies are characterized by the homotopy group π_2, which classifies the different mappings of the order parameter on the surface of a sphere S^2. The mappings of the surface of the sphere onto the interior of the sphere representing rotations $SO(3)$ of three-dimensional axes, for the combination of orbital order parameter and phase in the A phase, are all trivial. The only nontrivial topology arises when the sphere surrounding the point defect is mapped onto the sphere that represents the orientation of the spin axis $\mathbf{d}$ in the A phase. Such a mapping can be classified in terms of the winding number defined in equation (1.3), which describes how many times one sphere is wrapped around the other.

Boundary conditions can stabilize defects and textures which are unstable in a uniform medium. One obvious example of this is that ^{4}He in a rotating container has a vortex array as its equilibrium state. Similarly in the A phase of ^{3}He in a rotating container the equilibrium state either has an array of single vortices, or a texture made up of doubly quantized vortices like the one shown in Figure 7. The singly quantized vortices have a higher core energy than the texture, but the doubling of the circulation increases the energy contribution from regions a little further from the center of the texture.

A particularly interesting case of a texture imposed on the A phase by boundary conditions was discussed by Mermin [159]. In the A phase the orbital order parameter $\mathbf{l}$ tends to be lined up normal to the walls of the container. However, if one considers a simple container with the topology of a sphere there is no continuous way of arranging the axes $\mathbf{e}^{(1)}$, $\mathbf{e}^{(2)}$ so that they are always parallel to the surface of the sphere. Singularities of some sort have to be introduced on the surface. If there is a single singularity on the surface of the sphere, the rest of the surface, apart from

the immediate neighborhood of the singular point, can be opened up into a disk, and constant directions of $\mathbf{e}^{(1)}$, $\mathbf{e}^{(2)}$ corresponds to a rotation of these axes by 4π about the singularity. Alternatively one can have two singularities, opening up the sphere into something which is topologically equivalent to a finite cylinder, and parallel axes on this cylinder gives a phase change of 2π around each of the singularities, like the compass bearings we use on the earth's surface, with their singularities at each pole.

If there are two singly quantized vortices on the surface these must extend into the interior of the container. The two surface singularities must be connected by a line singularity. These two singularities on the surface can approach one another and merge, to form a single doubly quantized vortex on the surface, with no singular lines in the interior, only a texture which is known as a *boojum* [159]. Close to the surface the axes $\mathbf{e}^{(1)}$ and $\mathbf{e}^{(2)}$ rotate through an angle 4π, as they do in the outer parts of the texture shown in Figure 7. Further away from the surface the orientation of the vectors will change round a loop in the kind of way that they do on one of the inner loops round the center of this figure.

7   Liquid crystals

7.1   Order in liquid crystals

Liquid crystals, or mesophases, have some of the properties of solids and some of liquids. Typically they are anisotropic in space, but lack the rigidity of crystals, and can flow more or less like liquids. The molecules of a liquid crystal are generally quite complicated, but a physicist usually thinks of them as inflexible rod-like or disk-like objects whose two ends are indistinguishable – this is not because the two ends of the molecule are in fact indistinguishable, but because, although the molecules are aligned with their axes parallel to a certain direction, they are randomly pointing in two opposite directions. A survey of the properties of liquid crystals can be found in the book of de Gennes and Prost [160].

The simplest liquid crystal phase is known as the *nematic* phase. In nematics there is a preferred direction for the orientation of the molecular axis, but the order is otherwise like that of a liquid, with no ordering in space. The order parameter of the nematic phase is known as a *director,* which is like a vector, but with the two opposite directions equivalent. The space in which a director lives is a projective 2-sphere, a sphere with an equivalence relation between any two diametrically opposite points. In the equilibrium state of a bulk nematic the director is aligned everywhere in the same direction, but boundary conditions may impose other conditions on the director, such as that it should make a fixed angle with the boundary. Such boundary conditions mean that there will in general be a space-dependence

of the director, and there will usually be some sort of surface or interior defects asociated with changes in the director.

For the nematic phase the free energy density depends on the rate of change of the director **n**, according to the formula

$$\mathcal{F} = \frac{1}{2}K_1(\nabla \cdot \mathbf{n})^2 + \frac{1}{2}K_2(\mathbf{n} \cdot \nabla \times \mathbf{n})^2 + \frac{1}{2}K_3[\mathbf{n} \times (\nabla \times \mathbf{n})]^2 , \qquad (7.1)$$

where the three Frank constants K_1, K_2, K_3 control the free energies of splay, twist and bend respectively.

Most nematics are uniaxial, but there have been recent experimental studies of biaxial nematics, and these are particularly interesting from the topological point of view. We can think of the molecules of these as like rectangular disks, which, under suitable conditions, get stacked in such a way that both the normal to the disk and the orientation of the disk within the plane are determined. Their orientation is therefore defined by a triplet of axes, but with equivalence between those orientations that differ by a π rotation about one of the three axes. The order parameter therefore lives in the space $SO(3)/D_2$, where D_2 is the dihedral group of π rotations about the coordinate axes.

In the equilibrium state of a *cholesteric* liquid crystal the molecules are aligned along a direction that varies periodically. There are planes in which the director is constant, lying in the plane, but the director rotates uniformly around the axis perpendicular to these planes, maintaining a $\pi/2$ angle with the axis. Thus the director can be written as

$$\mathbf{n}(\mathbf{r}) = \hat{\mathbf{i}}\cos(2\pi\mathbf{r} \cdot \hat{\mathbf{k}}/l_0) + \hat{\mathbf{j}}\sin(2\pi\mathbf{r} \cdot \hat{\mathbf{k}}/l_0) , \qquad (7.2)$$

where $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ form a triplet of mutually perpendicular unit vectors. There is therefore a finite period $l_0/2$ in space between planes where the molecules are similarly oriented; the factor $1/2$ arises because we are dealing with a director, not a vector. The cholesteric has a broken $U(1)$ symmetry associated with this changing angle of orientation, as well as the director order parameter. From a practical point of view the cholesterics are particularly useful, as the periods of rotation can be comparable with optical wavelengths and can give dramatic optical interference effects which can be controlled by external electric fields or varying temperatures.

Blue phases of cholesteric liquid crystals appear to have a three-dimensional ordering of the director, generally in a cubic lattice.

Smectic phases of liquid crystals have some sort of spatial order as well as the orientational order of nematics. In the two phases I discuss here, smectic A and smectic C, the molecules lie on planes with a regular spacing between them, but there is no regular ordering of the molecules within the planes. The molecules therefore have one-dimensional ordering but no ordering in the other two dimensions. In smectic A the director along which

the molecules are preferentially aligned is normal to the planes on which the molecules lie, but in the smectic C phase the molecules are tilted relative to the normal to the planes in which the molecules lie.

There are also *columnar* phases of liquid crystals, in which there is spatial order in two dimensions, with the disk-like molecules stacked in columns. There is no long-range order within the columns, but the columns are arranged in a regular two-dimensional array.

7.2   Defects and textures

There are many similarities between the topological theory of defects in liquid crystals and the theory of defects in superfluid ^{3}He, which was discussed in Section 6. These theories were developed in parallel about twenty years ago in a series of papers by Toulouse *et al.* [132, 150, 161] and by Volovik and Mineev [151, 162]. A review of the topological study of defects in liquid crystals has been written by Kléman [163], and there is a more recent review by Kurik and Lavrentovich [164]. Defects in liquid crystals can be seen with a polarizing microscope, and so they have been studied for a long time. In fact the nematics get their name from the worm-like structures which can be seen threading them – these are simply disclinations.

The director, the order parameter of uniaxial nematic liquid crystals, lies on the surface of a projective sphere, a 2-sphere on which opposite points are regarded as identical. The only closed paths that cannot be shrunk to a point in this space are those paths that join two opposite points of the sphere. The homotopy group π_1 of this space is therefore the group Z_2 with two elements. There is one single type of line defect in a uniaxial nematic, which is a disclination around which the director rotates by an angle π. A disclination of this sort was shown for the spin of the A phase of superfluid ^{3}He in Figure 8.

The energy of a disclination in a nematic liquid crystal is much smaller than the energy of a disclination in a solid, because there is no regular arrangement of molecules to be disturbed by a strain field. Instead, equation (7.1) shows that the energy density depends on fields that fall off at least as fast as the reciprocal of the distance from the disclination core, so the energy of a disclination in a nematic liquid crystal, like the energy of a dislocation in a solid, is proportional to the length of the system in the direction of the disclination times the logarithm of the width of the system.

The point defects for a uniaxial nematic are defined by the mappings of a sphere surrounding the defect onto the projective sphere that represents the director. The situation is rather similar to that discussed for the A phase of superfluid ^{3}He discussed in Section 6.2, and the only difference that the sign ambiguity of the director makes is that the quantum numbers are ambiguous in sign. The quantum number is given again by equation (1.3), apart from this sign ambiguity, and the simplest form of defect is

again the hedgehog, with quantum number unity, with, for example, the director normal to the surface of the sphere at all points. The energy of a hedgehog is proportional to the radius of the system, as can be seen from equation (7.1), since the divergence of $\mathbf{n}$ is $2/r$ for this simple configuration of the hedgehog.

The order parameter for biaxial nematics defines a set of axes in space, as does the order parameter for the A phase of superfluid ^{3}He, but with the vital difference that rotation by π about any of the axes produces an equivalent orientation of the order parameter. This was discussed in by Poénaru and Toulouse [165]. This is particularly interesting, as it turns out that the residual symmetry of the order parameter is described by a non-commutative group, and the noncommutativity has some important consequences. In the previous subsection I referred to the space of this order parameter as $SO(3)/D_2$, but actually the topology of paths in $SO(3)$ is better described in terms of the covering group $SU(2)$, since $SO(3)$ does not distinguish rotation by 2π about an axis from no rotation, yet they are topologically distinct. There are equivalences of the order parameter given by the subgroup which is generated by the π rotations about the axes of the molecular ordering. This is the quaternion group Q, which is isomorphic to the eight elements $e, -e, \pm i\sigma_x, \pm i\sigma_y, \pm i\sigma_z$, where the σs are the Pauli matrices and e is the 2×2 identity matrix. Because this is a nonabelian group, with noncommuting elements, the defects behave in a different way from those we have discussed earlier. The space of the order parameter is $SU(2)/Q$, and the topologically distinct closed paths in this space are those that go from $\mathbf{e}$ to each of the five classes of this group. Rotations by $\pm\pi$ about the same axis are in the same class, because they are related by unitary transformations such as

$$-i\sigma_z = (i\sigma_x)(i\sigma_z)(-i\sigma_x) \, . \tag{7.3}$$

Therefore a path from the identity element $\mathbf{e}$ to $i\sigma_z$ is not distinct from a path from $\mathbf{e}$ to $-i\sigma_z$, but it is distinct from paths from $\mathbf{e}$ to $i\sigma_x$ or from $\mathbf{e}$ to $i\sigma_y$.

The linear defects for a biaxial nematic are therefore of four sorts, corresponding to the four classes of the quaternion group apart from the identity. Three are disclinations of strength $1/2$, corresponding to a π rotation about each of the three symmetry axes of the molecules, and one is a disclination of strength 1 corresponding to a 2π rotation about any axis. One possible way the molecular orientation can change on a loop round the defect is shown in Figure 9. These may look quite different if the molecules are differently oriented relative to the defect, but the different appearances can be continuously transformed into one another. For example, if each of the molecules shown in the figure is rotated by π about an axis through its
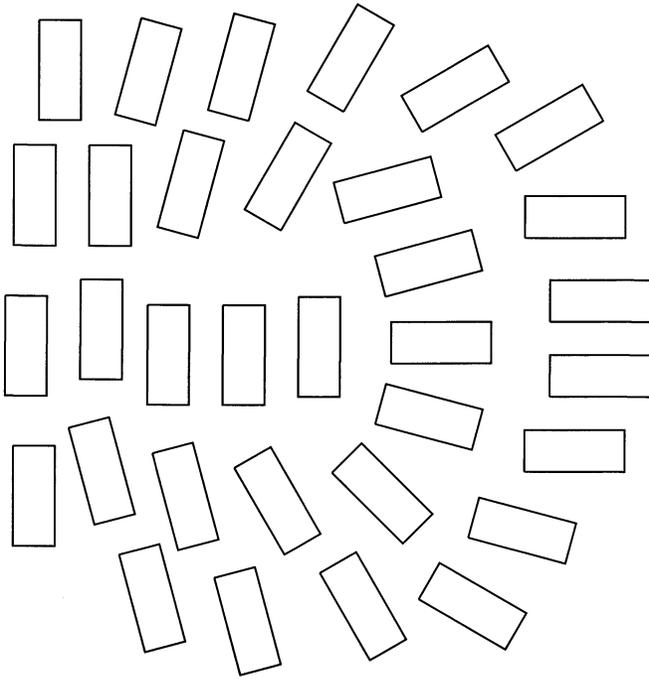
Fig. 9. Possible form for a disclination in which the orientation of molecules in a biaxial nematic rotates by π about the normal to the plane of the molecule.

centre in a fixed direction in the plane of the figure, the sense of rotation about the molecular axis is reversed.

The law of combination of such defects is given by the class multiplication law of the group. Two disclinations from the class $\{\pm\sigma_z\}$ can combine together either to annihilate one another, giving the group element e, or can combine to give a disclination of strength 1, corresponding to the group element $-e$. The same pair of defects can therefore be combined together in two completely different ways. One disclination from the class $\{\pm\sigma_x\}$ and one from the class $\{\pm\sigma_y\}$ will combine to give one from the class $\{\pm\sigma_z\}$. Another result of this special behavior of disclinations in biaxial nematics was pointed out by Toulouse [166]. Two different π disclinations cannot cross one another without generating a linking 2π disclination, as a careful examination of the changes generated by going round loops will show, so the disclinations give biaxial nematics a topological stiffness.

The topology of cholesterics is similar to that of biaxial nematics. The director at a given point is determined, in equation (7.2), by the triplet of vectors $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$, and this is invariant under sign reversal of any two of the
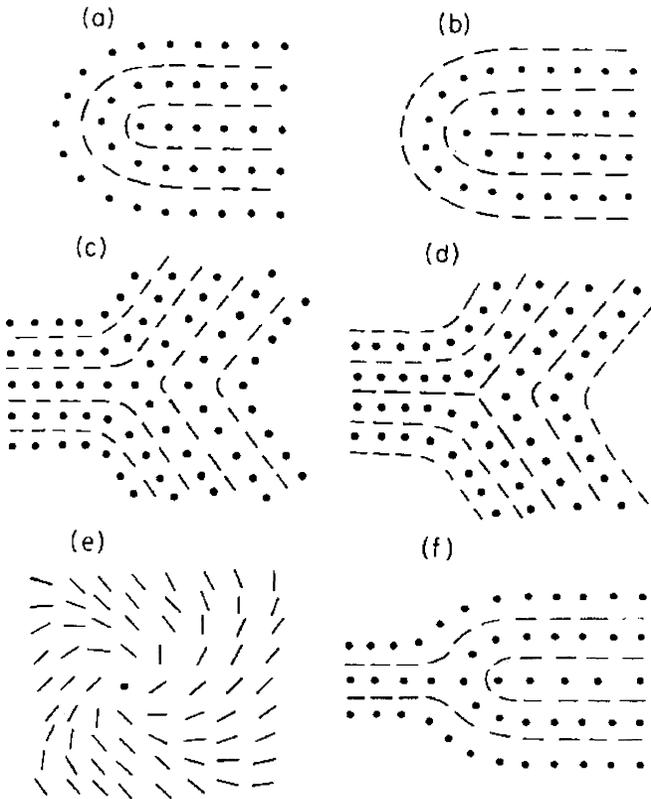
Fig. 10. Disclinations in a cholesteric liquid crystal. The dashes show curves in which the director is in the plane, while the dots show where the director is normal to the plane; in between the director is at an intermediate angle. A λ^+ is shown in (a), a τ^+ in (b), a λ^- in (c) and a τ^- in (d). One particular plane normal to the $\hat{\mathbf{k}}$ axis has the director oriented as shown in (e) for the χ disclination, and the director rotates steadily as one moves out of the plane, so that the surfaces of constant direction form a screw dislocation. In (f) a 2π disclination dissociated into two λ disclinations is shown.

three vectors, so the space in which the order parameter lives is $SO(3)/D_2$, for which the covering space describing the topology is $SU(2)/Q$. There are again four topologically distinct types of line defects corresponding to the four classes of the quaternion group Q distinct from the identity. Disclinations in a cholesteric liquid crystal were analysed by Kléman and Friedel [167] in terms of Volterra processes, and their classification is widely

(a)   **(b)**
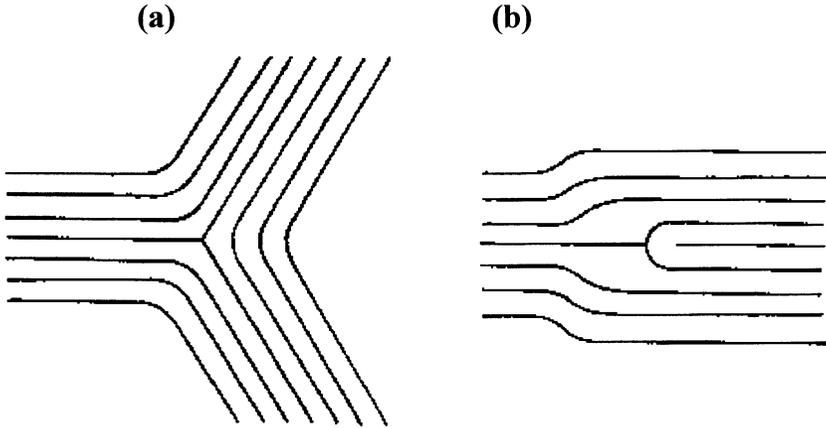
Fig. 11. A π disclination for a smectic A is shown in (a). When a disclination of opposite sign is added to this with a displacement of one lattice vector from the original disclination, the dislocation shown in (b) is obtained.

used. The disclinations τ^+, λ^+ are made by removing all the material beyond a certain plane normal to the $\hat{\mathbf{k}}$ plane (so that the director in that plane is everywhere in the same direction, which we identify as the $\hat{\mathbf{x}}$ direction). If the remaining material is folded about an axis parallel to the director so that the exposed surfaces are rejoined smoothly, a λ^+ disclination is obtained, as shown in Figure 10a. In this case the director is continuous even at the core of the disclination. If the remaining material is folded about an axis perpendicular to the director and the exposed surfaces are rejoined smoothly the τ^+ disclination shown in Figure 10b is obtained. In this case the director has singular behavior at the disclination core, and, as a result, the τ disclination core has a higher energy than the λ disclination core. The λ^-, τ^- disclinations can be constructed by the Volterra process in which a semi-infinite cut is made in a plane of constant director. If the cut terminates on an axis parallel to the director the λ^- disclination is obtained, but if it terminates on a line perpendicular to the director a τ^- disclination is obtained. The two faces exposed by the cut are then opened up by a π rotation until they form a plane, the remaining space is filled smoothly with undeformed material, then the system is allowed to relax until the angle is $2\pi/3$. The resulting disclinations are shown in Figures 10c and d. Again, the director is continuous at the core for the λ disclination and discontinuous at the core for the τ disclination. This is a manifestation of the fact that the $\lambda^\pm$ disclinations both belong to the class $\{\pm i\sigma_x\}$ of the homotopy group π_1 (a rotation of the pattern about the direction of the director), while the $\tau^\pm$ disclinations belong to the class $\{\pm i\sigma_y\}$

(rotation about an axis in the plane of constant director perpendicular to the director). Different patterns may be obtained if the disclination bends, or if there are other disclinations in its neighborhood, but the assignment to the class of this group is invariant.

The χ disclination could be formed by making a cut normal to the planes of constant phase and introducing a twist π of the director in any circle that goes round the core of the disclination. One of these planes, originally of constant director, is shown in Figure 10e. This defect could also be regarded as a screw dislocation with Burgers vector $l_0/2$, since surfaces of constant direction of the director have the form that one gets by introducing a screw dislocation into a set of parallel planes. It belongs to the class $\{\pm i\sigma_z\}$, since the rotation is about $\hat{\mathbf{k}}$, the normal to the planes of constant director. Like the τ disclination this has singular behavior of the director at the core. In principle it could dissociate into a $\lambda - \tau$ pair of disclinations, and the direction of the π change in angle can be reversed by bends in the line or the presence of other defects in the neighborhood.

Figure 10f shows a defect belonging to the class $\{-e\}$, a 2π disclination, in this case dissociated into a $\lambda^+ - \lambda^-$ pair, so that there is no singular behavior of the director at the core. This pair of disclinations half a period apart forms the same pattern that one would get with an edge dislocation, an extra period of the pattern inserted to the right of the position of the λ^-.

Neither biaxial nematics nor cholesterics have topologically stable point defects, since the mappings of a sphere onto $SO(3)$ are all trivial.

In the smectic A phase the direction of the molecules and the normal to the planes on which the molecules lie are the same, so the director has similar properties to the director of a uniaxial nematic. In the smectic C phase there is a pair of directions defined by the normal to the plane and the direction of the molecules. This is unchanged by a π rotation about the normal to the two, so the space of the directional order parameter is $SU(2)/Z_2$. Because the molecules are arranged on approximately planar surfaces there is a complicated interplay between directional order and positional order of these surfaces. Both disclinations and dislocations can exist.

In a regular solid the energy per unit length of a dislocation is proportional to the logarithm of the cross sectional area of the material, while the energy per unit length of a disclination is proportional to the cross-sectional area. In a smectic liquid crystal the positions of molecules can adjust to remove strain energy, but the spacings between surfaces must remain constant. As a consequence, the normal to one surface must also be normal to the neighboring surfaces. The splay $\nabla \cdot \mathbf{n}$ of a smectic A is given by the derivatives of the director in the plane of the surface. For a dislocation the splay falls off like $1/r^2$, while for a disclination it falls off like $1/r$, and so the energy per unit length of a dislocation is independent of the sample size, and the energy per unit length of a disclination depends logarithmically on

the cross-sectional area of the sample.

Figure 11a shows a π disclination in a smectic A liquid crystal. If a second disclination of opposite sign is formed by cutting away everything to the right of the middle of the surface of atoms immediately to the right of the disclination core, and then rejoining the two halves of this surface, the edge dislocation shown in Figure 11b is formed. Thus an edge dislocation can be regarded as the sum of two opposite disclinations displaced by a lattice spacing from one another. There is therefore an interplay between the algebras of dislocations and of disclinations.

The condition that the surfaces on which the molecules lie should have common normals, and, indeed, common centers of curvature, imposes severe restraints on the nature of point defects for A phase of a smectic. The only form of a true point defect is a hedgehog surrounded by spherical surfaces. More complicated structures may have terminating lines of defects. The C phase of a smectic cannot have isolated point defects, since there are no nontrivial mappings of a sphere onto the order parameter.

References

[1] Onsager L., *Nuovo Cimento* **6** (1949) 249-250.

[2] Vinen W.F., The detection of single quanta of circulation in liquid helium II, *Proc. Roy. Soc. (London) A* **260** (1961) 218-236.

[3] Rayfield G.W. and Reif F., Evidence for the creation and motion of quantized vortex rings in superfluid helium, *Phys. Rev. Lett.* **11** (1963) 305-308.

[4] Dirac PAM, Quantized singularities in the electromagnetic field, *Proc. Roy. Soc. (London)* **133** (1931) 60-72.

[5] Dirac PAM, The theory of magnetic poles, *Phys. Rev.* **74** (1948) 817-830.

[6] Piccard A. and Kessler E., Determination of the ratio between the electrostatic charges of the proton and of the electron, *Arch. Sci. Phys. Nat.* **7** (1925) 340-342.

[7] Petley B.W., The fundamental physical constants and the frontier of measurement (A. Hilger, Bristol, 1985) pp. 282-287.

[8] Donnelly R.J., in The collected works of Lars Onsager: with commentary (World Scientific, Singapore, 1996) pp. 693-696.

[9] London F., Superfluids, Vol. II, (1954) pp. 151-155.

[10] Thuneberg E.V., Introduction to the vortex sheet of superfluid ^{3}He, *Physica B* **210** (1995) 287-299.

[11] Parts Ü., Ruutu V.M.H., Koivuniemi J.H., Krusius M., Thuneberg E.V. and Volovik G.E., Measurements on the vortex sheet in rotating superfluid ^{3}He-A, *Physica B* **210** (1995) 311-333.

[12] Pitaevskii L.P., Vortex lines in an imperfect Bose gas, *Zhur. Eksp. Teor. Fiz.* **40** (1961) 454-477 [Translation in *Soviet Physics JETP* **13** (1961) 451].

[13] Gross E.P., Structure of a quantized vortex in boson systems, *Nuovo Cimento* **20** (1961) 454-477.

[14] Hall H.E. and Vinen W.F., The rotation of liquid helium II. I: Experiments on the propagation of second sound in uniformly rotating helium II, *Proc. Roy. Soc. (London) A* **238** (1956) 204.

[15] Hall H.E. and Vinen W.F., The rotation of liquid helium II. II: The theory of mutual friction in uniformly rotating helium II, *Proc. Roy. Soc. (London) A* **238** (1956) 215.

[16] Vinen W.F., Critical velocities in liquid helium II, in *Proceedings of the International School of Physics Enrico Fermi Course XXI. Liquid Helium,* edited by G. Careri (Academic Press, New York 1963) pp. 336-355.

[17] Langer J.S. and Fisher M.E., Intrinsic critical velocity of a superfluid, *Phys. Rev. Lett.* **19** (1967) 560-563.

[18] Muirhead C.M., Vinen W.F. and Donnelly R.J., The nucleation of vorticity by ions in rotating superfluid ^{4}He, *Phil. Trans. R. Soc. A* **311** (1984) 433-467.

[19] Donnelly R.J., Quantized vortices in helium II (Cambridge University Press, 1991).

[20] Feynman R.P., Application of quantum mechanics to liquid helium, in Progress in Low Temperature Physics 1, edited by C.J. Gorter (North-Holland, Amsterdam, 1955), pp. 17-53.

[21] Bardeen J., Cooper L.N. and Schrieffer J.R., Theory of superconductivity, *Phys. Rev.* **108** (1957) 1175-1204.

[22] London F., On the problem of the molecular theory of superconductivity, *Phys. Rev.* **74** (1948) 562-573.

[23] Putterman S.J., Superfluid Hydrodynamics (North-Holland, Amesterdam, 1974) pp. 404-407.

[24] Sonin E.B., Magnus force in superfluids and superconductors, *Phys. Rev. B* **55** (1997) 485-501.

[25] Whitmore S.C. and Zimmermann W., Observation of quantized circulation in superfluid helium, *Phys. Rev.* **166** (1968) 181-196.

[26] Zieve R.J., Close J.D., Davis J.C. and Packard R.E., New experiments on quantized circulation in superfluid ^{4}He, *J. Low Temp. Phys.* **90** (1993) 243-268.

[27] Zieve R.J., Mukharsky Y.M., Close J.D., Davis J.C. and Packard R.E., Investigation of quantized circulation in superfluid ^{3}He-B, *J. Low Temp. Phys.* **91** (1993) 315-339.

[28] Rayfield G.W. and Reif F., Quantized vortex rings in superfluid helium, *Phys. Rev.* **136** (1964) A1194-1208.

[29] Volovik G.E., Quantum-mechanical formation of vortices in a superfluid liquid, *Zhur. Eksp. Teor. Fiz. Pizma* **15** (1972) 116-120 [translation in *JETP Lett.* **15** (1972) 81-83].

[30] Deaver B.S. and Fairbank W.M., Experimental evidence for quantized flux in superconducting cylinders, *Phys. Rev. Lett.* **7** (1961) 43-46.

[31] Doll R. and Näbauer M., Experimental proof of magnetic flux quantization in a superconducting ring, *Phys. Rev. Lett.* **7** (1961) 51-52.

[32] Parks R.D. and Little W.A., Fluxoid quantization in a multiply-connected superconductor, *Phys. Rev.* **133** (1964) A97-102.

[33] Gough C.E., Colcough M.S., Forgan E.M., Jordan R.G., Keene M., Muirhead C.M., Rae A.I.M., Thomas N., Abell J.S. and Sutton S., Flux quantization in a high-T_c superconductor, *Nature (London)* **326** (1987) 855.

[34] Abrikosov A.A., On the magnetic properties of superconductors of the second type, *Zhur. Eksp. Teor. Fiz.* **32** (1957) 1442-1452.; *Sov. Phys. JETP* **5** (1957) 1174.

[35] Cribier D., Jacrot B., Madhov Rao L. and Farnoux B., Evidence from neutron diffraction for a periodic structure of the magnetic field in a niobium superconductor, *Phys. Lett.* **9** (1964) 106-107.

[36] Essmann U. and Träuble H. The direct observation of individual flux lines in type II superconductors, *Phys. Lett.* **24** (1967) 526-527.

[37] Cubitt R., Forgan E.M., Yang G., Lee S.L., Paul D.Mc.K., Mook H.A., Yethiraj M., Kes P.H., Li T.W., Menovsky A.A., Tarnawski Z. and Mortensen K., Direct observation of magnetic flux lattice melting and decomposition in the high-T_c superconductor $Bi_{2.15}Sr_{1.95}CaCu_2O_{8+x}$, *Nature* **365** (1993) 407-411.

[38] Bishop D.J. Gammel P.L., Huse D.A. and Murray C.A., Magnetic flux-line lattices and vortices in the copper oxide superconductors, *Science* **255** (1992) 165-172.

[39] Yarmchuk E.J., Gordon M.J.V. and Packard R.E., Observation of stationary vortex arrays in rotating superfluid helium, *Phys. Rev. Lett.* **43** (1979) 214-217.

[40] Yarmchuk E.J. and Packard R.E., Photographic studies of quantized vortex lines, *J. Low Temp. Phys.* **46** (1982) 479-515.

[41] Josephson B.D., Possible new effects in superconductive tunnelling, *Phys. Lett.* **1** (1962) 251-253.

[42] Josephson B.D., Supercurrents through barriers, *Adv. Phys.* **14** (1965) 419-451.

[43] Anderson P.W., The Josephson effect and quantum coherence measurements in superconductors and superfluids, in Progress in Low Temperature Physics 5, edited by C.J. Gorter (North-Holland, Amsterdam, 1967) pp. 1-43.

[44] Parker W.H., Taylor B.N. and Langenberg D.N., Measurement of 2e/h using the ac Josephson effect and its implications for quantum electrodynamics, *Phys. Rev. Lett.* **18** (1967) 287-291.

[45] Cohen E.R. and Taylor B.N., The 1986 adjustment of the fundamental physical constants, *Rev. Mod. Phys.* **59** (1987) 1121.

[46] Clarke J., Experimental comparison on the Josephson voltage-frequency relation in different superconductors, *Phys. Rev. Lett.* **21** (1968) 1566-1569.

[47] Tsai J.S., Jain A.K. and Lukens E., High-precision test of the universality of the Josephson voltage-frequency relation, *Phys. Rev. Lett.* **51** (1983) 316-319.

[48] Kautz R.L. and Lloyd F.L., Precision of series-array Josephson voltage standards, *Appl. Phys. Lett.* **51** (1987) 2043-2045.

[49] Duan J.M., Mass of a vortex line in superfluid ^{4}He: effects of gauge-symmetry breaking, *Phys. Rev. B* **49** (1994) 12381-12383.

[50] Demircan E., Ao P. and Niu Q., Vortex dynamics in superfluids: cyclotron-type motion, *Phys. Rev. B* **54** (1996) 10027-10034.

[51] Thouless D.J., Ao P., Niu Q., Geller M.R. and Wexler C., Quantized vortices in superfluids and superconductors. In the Proceedings of the 9th International Conference on Recent Progress in Many-Body Theories, edited by D. Neilson and R.F. Bishop (World Scientific, Singapore, 1998) 387-398.

[52] Berry M.V., Quantal phase factors accompanying adiabatic changes, *Proc. Roy. Soc. (London) A* **392** (1984) 45-57.

[53] Haldane F.D.M. and Wu Y.S., Quantum dynamics and statistics of vortices in two-dimensional superfluids, *Phys. Rev. Lett.* **55** (1985) 2887-2890.

[54] Thouless D.J., Ao P. and Niu Q., Vortex dynamics in superfluids and the Berry phase, *Physica A* **200** (1993) 42-49.

[55] Thouless D.J., Ao P. and Niu Q., Transverse force on a quantized vortex in a superfluid, *Phys. Rev. Lett.* **76** (1996) 3758-3761.

[56] Tang J.M. and Thouless D.J., Longitudinal force on moving potential, *Phys. Rev. B* **58** (1998) 14179-2.

[57] Wexler C., Magnus and Iordanskii forces in superfluids, *Phys. Rev. Lett.* **79** (1997) 1321-1324.

[58] Laughlin R.B., Quantized Hall conductivity in two dimensions, *Phys. Rev. B* **23** (1981) 5632-5633.

[59] Volovik G.E., Comment on "Transverse force on a quantized vortex in a superfluid", *Phys. Rev. Lett.* **77** (1997) 4687.

[60] Iordanskii S.V., On the mutual friction between the normal and superfluid components in a rotating Bose gas, *Ann. Phys. (NY)* **29** (1964) 335-349; Iordanskii S.V., *Zhur. Eksp. Teor. Fiz.* **49** (1965) 225-236 [translation in *Soviet Phys. JETP* **22** (1965) 160-167].

[61] Geller M.R., Wexler C. and Thouless D.J., Transverse force on a quantized vortex in a superconductor, *Phys. Rev. B* **57** (1998) R8119-8122.

[62] Dirac P.A.M., Peierls R.E. and Pryce M.H.L., On Lorentz invariance in the quantum theory, *Proc. Cambridge Philos. Soc.* **44** (1942) 143-157.

[63] Nozières P. and Vinen W.F., *Phil. Mag.* **14** (1966) 667.

[64] Klitzing Kv., Dorda G. and Pepper M., New method for high-accuracy determination of fine-structure constant based on quantized Hall resistance, *Phys. Rev. Lett.* **45** (1980) 494-497.

[65] Yoshihiro K., Kinoshita J., Inayaki K., Yamanouchi C., Endo T., Murayama Y., Koyanagi M., Yagi A., Wakabayashi J. and Kawaji S., Quantum Hall effect in silicon metal-oxide-semiconductor inversion layers: Experimental conditions for determination of h/e^2, *Phys. Rev. B* **33** (1986) 6874-6896.

[66] Hartland A., Jones K., Williams J.M., Gallagher B.L. and Galloway T., Direct comparison of the quantized Hall resistance in gallium arsenide and silicon, *Phys. Rev. Lett.* **66** (1991) 969-973.

[67] Taylor B.N., New measurement standards for 1990, *Physics Today* **42** (1989) 23-26.

[68] Aoki H. and Ando T., Effect of localization on the Hall conductivity in the two-dimensional system in strong magnetic fields, *Solid St. Commun.* **38** (1981) 1079-1082.

[69] Prange R.E., Quantized Hall resistance and the measurement of the fine-structure constant, *Phys. Rev. B* **23** (1981) 4802-4805.

[70] Thouless D.J., Localization and the two-dimensional Hall effect, *J. Phys. C* **14** (1981) 3475-3480.

[71] Thouless D.J., Kohmoto M., Nightingale M.P. and den Nijs M., Quantized Hall conductance in a two-dimensional periodic potential, *Phys. Rev. Lett.* **49** (1982) 405-408.

[72] Streda P., Theory of quantised Hall conductivity in two dimensions, *J. Phys. C* **15** (1982) L717-721.

[73] Avron J.E. and Seiler R., Quantization of the Hall conductance of general multiparticle Schrödinger Hamiltonians, *Phys. Rev. Lett.* **54** (1985) 259-262.

[74] Niu Q., Thouless D.J. and Wu Y.S., Quantized Hall conductance as a topological invariant, *Phys. Rev. B* **31** (1985) 3372-3377.

[75] Büttiker M., Absence of backscattering in the quantum Hall effect in multiprobe conductors, *Phys. Rev. B* **38** (1988) 9375-9389.

[76] Tsui D.C., Stormer H.L. and Gossard A.C., Two-dimensional magneto-transport in the extreme quantum limit, *Phys. Rev. Lett.* **48** (1982) 1559-1562.

[77] Kubo R., Statistical-mechanical theory of irreversible processes. I. General theory and simple applications to magnetic and conduction problems, *J. Phys. Soc. Jpn.* **12** (1957) 570-586.

[78] Kubo R. The fluctuation-dissipation theorem, *Rep. Progr. Phys.* **29** (1966) 255-284.

[79] Lee P.A. and Ramakrishnan T.V., Dirordered electronic systems, *Rev. Mod. Phys.* **57** (1985) 287-337.

[80] Bohm D., Note on a theorem of Bloch concerning possible causes of superconductivity, *Phys. Rev.* **75** (1949) 502-504.

[81] Halperin B.I., Quantized Hall conductance, current-carrying edge states, and the existence of extended states in a two-dimensional disordered potential, *Phys. Rev. B* **25** (1982) 2185-2190.

[82] Thouless D.J., Edge voltages and distributed currents in the quantum Hall effect, *Phys. Rev. Lett.* **71** (1993) 1879-1882.

[83] Azbel M.Ya., Energy spectrum of conduction electrons in a magnetic field, *Zh. Eksp. Teor. Fiz.* **46** (1964) 929-947 [translation in *Soviet Phys. JETP* **19** (1964)].

[84] Hofstadter D., Energy levels and wave functions of Bloch electrons in rational and irrational magnetic fields, *Phys. Rev. B* **14** (1976) 2239-2249.

[85] Niu Q. and Thouless D.J., Quantum Hall effect with realistic boundary conditions, *Phys. Rev. B* **35** (1987) 2188-2197.

[86] Choquet-Bruhat Y., DeWitt-Morette C. and Dillard-Bleick M., Analysis, Manifolds and Physics (North-Holland Publishing Co., Amsterdam 1982), pp. 393-396.

[87] Thouless D.J., Wannier functions for magnetic sub-bands, *J. Phys. C* **17** (1984) L325-327.

[88] Kohmoto M., Topological invariant and the quantization of the Hall conductance, *Ann. Phys. (NY)* **160** (1985) 343-354.

[89] Arovas D.P., Bhatt R.N., Haldane F.D.M., Littlewood P.B. and Rammal R., Localization, wave-function topology, and the integer quantized Hall effect, *Phys. Rev. Lett.* **60** (1988) 619-622.

[90] Laughlin R.B., Anomalous quantum Hall effect: an incompressible quantum fluid with fractionally charged excitations, *Phys. Rev. Lett.* **50** (1983) 1395-1398.

[91] Haldane F.D.M. and Rezayi E.H., Finite-size studies of the incompressible state of the fractionally quantized Hall effect and its excitations, *Phys. Rev. Lett.* **54** (1985) 237-240.

[92] Jain J.K., Composite fermion approach for the fractional quantum Hall effect, *Phys. Rev. Lett.* **63** (1989) 199-202.

[93] Jain J.K., Microscopic theory of the fractional quantum Hall effect, *Adv. Phys.* **41** (1992) 105-146.

[94] Jain J.K. and Kamilla R.K., Composite fermions in the Hilbert space of the lowest electronic Landau level, *Int. J. Mod. Phys. B* **11** (1997) 2621-2660.

[95] Anderson P.W., Remarks on the Laughlin theory of the fractionally quantized Hall effect, *Phys. Rev. B* **28** (1983) 2264-2265.

[96] Tao R. and Wu Y.S., Gauge invariance and the fractional quantum Hall effect, *Phys. Rev. B* **30** (1984) 1097-1098.

[97] Thouless D.J., Level crossing and the fractional quantum Hall effect, *Phys. Rev. B* **40** (1989) 12034-12036.

[98] Thouless D.J. and Gefen Y., Fractional quantum Hall effect and multiple Aharonov-Bohm periods, *Phys. Rev. Lett.* **66** (1991) 806-909.

[99] Gefen Y. and Thouless D.J., Detection of fractional charge and quenching of the quantum Hall effect, *Phys. Rev. B* **47** (1993) 10423-104236.

[100] Halperin B.I., Theory of the quantized Hall conductance, *Helv. Phys. Acta* **56** (1983) 75-102.

[101] Sondhi S.L., Karlhede A., Kivelson S.A. and Rezayi E.H., Skyrmions and the crossover from the integer to fractional quantum Hall effect at small Zeeman energies, *Phys. Rev. B* **47** (1993) 16419-16426.

[102] Fertig H.A., Brey L., Cote R. and MacDonald A.H., Charged spin-texture excitations and the Hartree-Fock approximation in the quantum Hall effect, *Phys. Rev. B* **50** (1994) 11018-11021.

[103] Xie X.C. and He S., Skyrmion excitations in quantum Hall systems, *Phys. Rev. B* **53** (1996) 1046-1049.

[104] MacDonald A.H., Fertig H.A. and Brey L., Skyrmions without sigma models in quantum Hall ferromagnets, *Phys. Rev. Lett.* **76** (1996) 2153-2156.

[105] Fertig H.A., Brey L., Côté R., Macdonald A.H., Karlhede A. and Sondhi S.L., Hartree-Fock theory of skyrmions in quantum Hall ferromagnets, *Phys. Rev.* **55** (1997) 10671-10680.

[106] Barrett S.E., Dabbagh G., Pfeiffer L.N. and West K.W., Optically pumped NMR evidence for finite-size Skyrmions in GaAs quantum wells near Landau level filling $\nu = 1$, *Phys. Rev. Lett.* **74** (1995) 5112-5115.

[107] Schmeller A., Eisenstein J.P., Pfeiffer L.N. and West K.W., Evidence for Skyrmions and single spin flips in the integer quantized Hall effect, *Phys. Rev. Lett.* **75** (1995) 4290-4293.

[108] Aifer E.H., Goldberg B.B. and Broido D.A., Evidence of Skyrmion excitations about $\nu = 1$ in n-modulation-doped single quantum wells by interband optical transmission, *Phys. Rev. Lett.* **76** (1996) 680-683.

[109] Bayot V., Grivei E., Melinte S., Santos M.B. and Shayegan M., Giant low temperature heat capacity of GaAs quantum wells near Landau level filling $\nu = 1$, *Phys. Rev. Lett.* **76** (1996) 4584-4587.

[110] Maude D.K., Potemski M., Portal J.C., Henini M., Eaves L., Hill G. and Pate M.A., Spin excitations of a two-dimensional electron gas in the limit of vanishing Lande g factor, *Phys. Rev. Lett.* **77** (1996) 4604-4607.

[111] Peierls R.E., Remarks on transition temperatures, *Helv. Phys. Acta* **7** (1934) 81-83. Peierls R.E., Some properties of solids, *Ann. Inst. Henri Poincaré* **5** (1935) 177-222.

[112] Mermin N.D. and Wagner H., Absence of ferromagnetism in one- and two-dimensional isotropic Heisenberg models, *Phys. Rev. Lett.* **17** (1966) 1133.

[113] Hohenberg P.C., Existence of long-range order in one and two dimensions, *Phys. Rev.* **158** (1967) 383.

[114] Imry Y. and Gunther L., Fluctuations and physical properties of the two-dimensional crystal lattice, *Phys. Rev. B* **3** (1971) 3939-3945.

[115] Berezinskii V.L., Destruction of long-range order in one-dimensional and two-dimensional systems with a continuous symmetry group. I. Classical systems, *Zhur. Eksp. Teor. Fiz.* **59** (1970) 907 [translation in *Sov. Phys. JETP* **32** (1970) 493].

[116] Berezinskii V.L., Destruction of long-range order in one-dimensional and two-dimensional systems with a continuous symmetry group. II. Quantum systems, *Zhur. Eksp. Teor. Fiz.* **61** (1971) 1144.

[117] Kosterlitz J.M. and Thouless D.J., Long range order and metastability in two dimensional solids and superfluids, *J. Phys. C* **5** (1972) L124-126.

[118] Kosterlitz J.M. and Thouless D.J., Ordering, metastability and phase transitions in two-dimensional systems, *J. Phys. C* **6** (1973) 1181-1203.

[119] Wegner F.J., Spin-ordering in a planar classical Heisenberg model, *Z. Phys.* **206** (1967) 465-470.

[120] Jancovici B., Infinite susceptibility without long-range order: the two-dimensional harmonic "solid", *Phys. Rev. Lett.* **19** (1967) 20-22.

[121] Nelson D.R. and Kosterlitz J.M., Universal jump in the superfluid density of two-dimensional superfluids, *Phys. Rev. Lett.* **39** (1977) 1201-1205.

[122] Anderson P.W. and Yuval G., Some numerical results on the Kondo problem and the inverse square one-dimensional Ising model, *J. Phys. C* **4** (1971) 607-620.

[123] Kosterlitz J.M., The critical properties of the two-dimensional xy model, *J. Phys. C* **7** (1974) 1046-1060.

[124] José J., Kadanoff L.P., Kirkpatrick S. and Nelson D.R., Renormalization, vortices and symmetry breaking perturbations in the two-dimensional planar model, *Phys. Rev. B* **16** (1977) 1217-1241.

[125] Nelson D.R., Defect-mediated phase transitions, in Phase transitions and critical phenomena Vol. 7, pp. 1-99, edited by C. Domb and J.L. Lebowitz (Academic Press Ltd., London and New York, 1983).

[126] Rudnick I., Critical surface density of the superfluid component in ^{4}He films, *Phys. Rev. Lett.* **40** (1978) 1454-1455.

[127] Bishop D.J. and Reppy J.D., Study of the superfluid transition in two-dimensional ^{4}He films, *Phys. Rev. Lett.* **40** (1978) 1727-1730.

[128] Bishop D.J. and Reppy J.D., Study of the superfluid transition in two-dimensional ^{4}He films, *Phys. Rev. B* **22** (1979) 5171-5185.

[129] Ambegaokar V., Halperin B.I., Nelson D.R. and Siggia E.D., Dynamics of superfluid films, *Phys. Rev. B* **21** (1980) 1806-1826.

[130] McQueeney D., Agnolet G.T. and Reppy J.D., Surface superfluidity in dilute ^{4}He–^{3}He mixtures, *Phys. Rev. Lett.* **52** (1984) 1325-1328.

[131] Polyakov A.M., Interaction of Goldstone particles in two dimensions. Applications to ferromagnets and massive Yang-Mills fields, *Phys. Lett. B* **59** (1975) 79-81.

[132] Kléman M., Relationship between Burgers circuit, Volterra process and homotopy groups, *J. Phys. Lett. (Paris)* **38** (1977) L199-202.

[133] Nabarro F.R.N., Theory of crystal dislocations (Clarendon Press, Oxford, 1967).

[134] Young A.P., Melting and the vector Coulomb gas in two dimensions, *Phys. Rev. B* **19** (1979) 1855-1866.

[135] Halperin B.I. and Nelson D.R., Theory of two-dimensional melting, *Phys. Rev. Lett.* **41** (1978) 121-124 and 1519.

[136] Nelson D.R. and Halperin B.I., Dislocation-mediated melting in two dimensions, *Phys. Rev. B* **19** (1979) 2457-2484.

[137] Grimes C.C. and Adams G., Evidence for a liquid-to-crystal phase transition in a classical, two-dimensional sheet of electrons, *Phys. Rev. Lett.* **42** (1979) 795-798.

[138] Gallet F., Deville G., Valdes A. and Williams F.I.B., Fluctuations and shear modulus of a classical two-dimensional electron solid: experiment, *Phys. Rev. Lett.* **49** (1982) 212-215.

[139] Morf R.H., Temperature dependence of the shear modulus and melting of the two-dimensional electron solid, *Phys. Rev. Lett.* **43** (1979) 931-935.

[140] Strandburg K.J., Two-dimensional melting, *Rev. Mod. Phys.* **60** (1988) 161-207.

[141] Beasley M.R., Mooij J.E. and Orlando T.P., Possibility of vortex–antivortex pair dissociation in two-dimensional superconductors, *Phys. Rev. Lett.* **42** (1979) 1165-1168.

[142] Doniach S. and Huberman B.A., Topological excitations in two-dimensional superconductors, *Phys. Rev. Lett.* **42** (1979) 1169-1172.

[143] Hebard A.F. and Fiory A.T., Evidence for the Kosterlitz-Thouless transition in thin superconducting aluminum films, *Phys. Rev. Lett.* **44** (1981) 291-294.

[144] Fiory A.T., Hebard A.F. and Glaberson W.I., Superconducting phase transitions in indium/indium-oxide thin-film composites, *Phys. Rev. B* **28** (1983) 5075-5087.

[145] Hebard A.F. and Fiory A.T., Critical-exponent measurements of a two-dimensional superconductor, *Phys. Rev. Lett.* **50** (1983) 1603-1606.

[146] Minnhagen P., The two-dimensional Coulomb gas, vortex unbinding, and superfluid–superconducting films, *Rev. Mod. Phys.* **59** (1987) 1001-1066.

[147] Huberman B.A. and Doniach S., Melting of two-dimensional vortex lattices, *Phys. Rev. Lett.* **43** (1979) 950-952.

[148] Fisher D.S., Flux-lattice melting in thin-film superconductors, *Phys. Rev. B* **22** (1980) 1190-1199.

[149] Abraham D.A., Lobb C.J., Tinkham M. and Klapwijk T.M., Resistive transition in two-dimensional arrays of superconducting weak links, *Phys. Rev. B* **26** (1982) 5268-5271.

[150] Toulouse G. and Kléman M., Principles of a classification of defects in disordered media, *J. Phys. Lett. France* **37** (1976) L149-151.

[151] Volovik G.E. and Mineev V.P., Investigation of singularities in superfluid He^3 and liquid crystals by homotopic topology methods, *Zhur. Eksp. Teor. Fiz.* **72** (1977) 2256-2274 [translation in *Soviet Phys. JETP* **45** (1977) 1186-1196].

[152] Volovik G.E., Exotic Properties of Superfluid ^{3}He (World Scientific, Singapore, 1992).

[153] Anderson P.W. and Toulouse G., Phase slippage without vortex cores: vortex textures in superfluid ^{3}He, *Phys. Rev. Lett.* **38** (1977) 508-511.

[154] Parts Ü., Avilov V.V., Koivuniemi J.H., Krusius M., Ruohio J.J. and Ruutu V.M.H., Vortex arrays of coexisting singly and doubly quantized vortex lines in ^{3}He-A, *Czechoslovak J. Phys.* **46** (1996) 13-14.

[155] Krusius M., The vortices of superfluid ^{3}He, *J. Low Temp. Phys.* **91** (1993) 233-273.

[156] Thuneberg E.V., Identification of vortices in superfluid ^{3}He-B, *Phys. Rev. Lett.* **56** (1986) 359-362.

[157] Salomaa M.M. and Volovik G.E., Topological transition of v-vortex core matter in ^{3}He-B, *Europhys. Lett.* **2** (1986) 781-787.

[158] Kondo Y., Korhonen J.S., Krusius M., Dmitriev V.V., Mukharsky Y.M., Sonin E.B. and Volovik G.E., Direct observation of the nonaxisymmetric vortex in superfluid ^{3}He-B, *Phys. Rev. Lett.* **67** (1991) 81-84.

[159] Mermin N.D., Surface singularities and superflow in ^{3}He–A, in Quantum Fluids and Solids, edited by S.B Trickey, E.D. Adams and J.W. Dufty (Plenum Press, New York, 1977), pp. 3-22.

[160] de Gennes P.G. and Prost J., The Physics of Liquid Crystals (Clarendon Press, Oxford, 1993).

[161] Kléman M. and Michel L., Spontaneous breaking of Euclidean invariance and classification of topologically stable defects and configurations of crystals and liquid crystals, *Phys. Rev. Lett.* **40** (1978) 1387-1390.

[162] Volovik G.E. and Mineev V.P., Line and point singularities in superfluid He^3, *Pis'ma Zh. Eksp. Teor. Fiz.* **24** (1976) 605-608 [translation in *JETP Lett.* **24** (1976) 561-563].

[163] Kléman M., Points, Lines and Walls (John Wiley & Sons, Chichester, 1983).

[164] Kurik M.V. and Lavrentovich O.D., Defects in liquid crystals: homotopy theory and experimental studies, *Usp. Fiz. Nauk* **154** (1988) 381-431 [translation in *Sov. Phys. Usp.* **31** (1988) 196-224].

[165] Poénaru V. and Toulouse G., The crossing of defects in ordered media and the topology of 3–manifolds, *J. Phys. (Paris)* **38** (1977) 887-895.

[166] Toulouse G., On biaxial nematics, *J. Phys. Lett. (Paris)* **38** (1977) L67-68.

[167] Kléman M. and Friedel J., *J. Phys. France Colloq* **30** (1969) 43.