Available online at www.sciencedirect.com
SCIENCE DIRECT®

Annals of Physics 319 (2005) 217–249

ANNALS of PHYSICS

www.elsevier.com/locate/aop

Review

Is electromagnetic gauge invariance spontaneously violated in superconductors?

Martin Greiter *

Institut für Theorie der Kondensierten Materie, Universität Karlsruhe, Postfach 6980, D-76128 Karlsruhe, Germany

Received 17 January 2005; accepted 29 March 2005
Available online 19 May 2005

Abstract

We aim to give a pedagogical introduction to those elementary aspects of superconductivity which are not treated in the classic textbooks. In particular, we emphasize that global $U(1)$ phase rotation symmetry, and not gauge symmetry, is spontaneously violated, and show that the BCS wave function is, contrary to claims in the literature, fully gauge invariant. We discuss the nature of the order parameter, the physical origin of the many degenerate states, and the relation between formulations of superconductivity with fixed particle numbers vs. well-defined phases. We motivate and to some extend derive the effective field theory at low temperatures, explore symmetries and conservation laws, and justify the classical nature of the theory. Most importantly, we show that the entire phenomenology of superconductivity essentially follows from the single assumption of a charged order parameter field. This phenomenology includes Anderson's characteristic equations of superfluidity, electric and magnetic screening, the Bernoulli Hall effect, the balance of the Lorentz force, as well as the quantum effects, in which Planck's constant manifests itself through the compactness of the $U(1)$ phase field. The latter effects include flux quantization, phase slippage, and the Josephson effect.

PACS: 74.20.−z; 11.15.−q; 11.15.Ex

Keywords: Superconductivity; Gauge invariance; Order parameter; Effective field theory; Higgs mechanism

* Fax: +49 721 608 7779.
E-mail address: greiter@tkm.uni-karlsruhe.de.

0003-4916/$ - see front matter
doi:10.1016/j.aop.2005.03.008

1. Introduction

Many years ago, Steven Weinberg mentioned to me that he was disconcerted that none of the classic textbooks on superconductivity would explain the phenomenon in terms of the Higgs mechanism [1] for the electromagnetic gauge field. This concern is of course very well justified, and it was most likely with this concern in mind that Weinberg has included a section on superconductivity in his treatment of spontaneous symmetry breaking and the Higgs mechanism in his series of volumes entitled *The quantum theory of fields* [2]. When I was asked recently to present a series of lectures on superconductivity, I opened his book expecting to find a particularly lucid exposition of this in condensed matter physics rarely emphasized perspective. I found the exposition I was looking for, but to my surprise, build around the following statement: *A superconductor is simply a material in which electromagnetic gauge invariance is spontaneously broken.* What Weinberg means with this statement is just that the electromagnetic gauge field "acquires a mass" due to the Higgs mechanism in a superconductor, as particle physicists often speak of spontaneously broken gauge invariance interchangeably with the Higgs mechanism.

Nonetheless, I am not perfectly at ease with the above statement, which is, by the way, by no means specific to Weinberg's exposition, but widely believed and accepted. While it is obvious that Weinberg fully understands the matter, the statement may still be misleading to a young student who is learning the subject for the first time. The problem is that the statement is, if one takes it literally, not correct: gauge invariance cannot spontaneously break down as a matter of principle, and in particular is not broken in a superconductor, as I will explain in the following section.

This paper is organized as follows. In Section 2, we discuss the statement quoted above including the danger which may result from a literal interpretation of it in depth. In particular, we show that the BCS ground state is, in contrast to statements made in the literature, fully gauge invariant. The crucial ingredient often omitted is that gauge transformations involve, in addition to the standard transformation of gauge fields, local phase rotations of *both* creation (and annihilation) operators *and* wave functions. In Section 3, we discuss the nature of the order parameter in superconductors, with particular emphasis on finite systems, which always possess a unique ground state. The arising subtlety is explained by drawing an analogy to quantum antiferromagnets, which also possess a unique and rotationally invariant ground state for finite systems. In Section 4, we motivate and elaborate the effective field theory of a superconductor at low temperatures, which contains the theory of a neutral superfluid as the special case where the charge is set to zero. In particular, we obtain the particle density and current as well as the energy and momentum density from the physical symmetries of the theory, invariance under global $U(1)$ phase rotations of the order parameter and invariance under translations in time and space. The quest for a consistent definition of the superfluid velocity yields a relation between current and momentum densities in the superfluid, which in turn requires corrections to the effective Lagrangian. Since the density of the superfluid is essentially the "momentum conjugate" of the order parameter phase, Hamilton's equations yield physical information not contained in the Euler–Lagrange equations; specifically, we obtain a gauge invariant generalization of Anderson's characteristic equations of superfluidity to the case of superconductors. We conclude this section with a brief justification of the classical nature of the effective field theory. In Section 5, we discuss the phenomenology of superconductors as compared to neutral superfluids, or, in general terms, the Higgs mechanism. To begin with, we briefly address the phenomenology of neutral superfluids including vortex quantization, and give a general introduction to the Higgs mechanism in field theories. We then turn to the phenomenology of simply connected superconductors, solve the equations of motion, obtain electric and magnetic screening, London's equation, the Bernoulli Hall effect, and the balance of the Lorentz force. We demonstrate that the Higgs mechanism never corresponds to a spontaneous violation of a gauge symmetry, and that it is incorrect to interpret it in terms of "a mass acquired by the electromagnetic gauge field," as the massive field is no longer a gauge field. Specifically, we show that the massive vector field, which may alternatively be used to describe a (simply connected) superconductor, is correctly interpreted as a four-vector formed by the chemical potential and the three components of the superfluid velocity. We conclude this section with a discussion of the subtle difference between the physical invariance of the theory under global $U(1)$ rotations of the order parameter phase and gauge invariance, which is nothing but a local invariance of our description of the system. In the last section, we review a family of "quantum effects": the quantization of magnetic flux in superconductors, phase slippage, and the Josephson effect in both neutral superfluids and superconductors. In these effects, Planck's constant manifests itself in the phenomenology through the compactness of the order parameter phase field; these effects require either a non-trivial topology or more than one superfluid. We derive them from the effective field theory introduced in Section 4, and thereby demonstrate that the very few assumptions made in motivating the effective theory are sufficient to account for them.

2. Gauge invariance

To illustrate how dangerous the statement quoted in the introduction is in the case of superconductivity, where we do not only have a description in terms of an effective field theory but also a microscopic description in terms of model Hamiltonians and trial wave functions, I will at first assume the statement was true and take it literally. I will pretend to be a student who has just learned that electromagnetic gauge invariance is spontaneously violated in a superconductor. Well, what does this mean? A spontaneously broken symmetry means that the Hamiltonian of a given system in the thermodynamic limit is invariant under a given symmetry transformation (i.e., commutes with the generator(s) of this symmetry) while the ground state is not invariant. There are many ground states, which transform into each other under the symmetry transformations. A classic example is ferromagnetism: The Hamiltonian is rotationally invariant, while any particular ground state, specified by the direction the magnetization vector points to, is not. So if gauge invariance is broken in a superconductor, this must mean that the ground state of the superconductor does not share the gauge invariance of the Hamiltonian. Indeed, a glance at the BCS wave function [3–6]

$$|\psi_\phi\rangle = \prod_{\boldsymbol{k}} (u_{\boldsymbol{k}} + v_{\boldsymbol{k}} e^{i\phi} c^\dagger_{\boldsymbol{k}\uparrow} c^\dagger_{-\boldsymbol{k}\downarrow})|0\rangle, \tag{1}$$

where the coefficients $u_{\boldsymbol{k}}$ and $v_{\boldsymbol{k}}$ are chosen real and ϕ is an arbitrary phase, appears to confirm this picture. There are many different ground state wave functions, labeled by ϕ, which transform into each other under an electromagnetic gauge transformation given by

$$c^\dagger_{\boldsymbol{k}\sigma} \to e^{i\frac{e}{\hbar c}\Lambda} c^\dagger_{\boldsymbol{k}\sigma}, \tag{2}$$

which is tantamount to taking

$$\phi \to \phi + \frac{2e}{\hbar c}\Lambda.$$

For simplicity, we have chosen Λ independent of spacetime. The electron charge throughout this article is $-e$.

Next, I the student ask myself whether these many BCS wave functions for different parameters ϕ correspond to physically different states. I know that gauge transformations are not physical transformations: gauge invariance is an invariance of a description of a system, while other symmetries correspond to invariances under physical transformations, like rotations or translations, which affect the physical state in question. For example, if the Hamiltonian for given system (like a ferromagnet) is invariant under rotations in space, this implies that if we rotate a given eigenstate, we will obtain another eigenstate. Depending on whether the original state is rotationally invariant or not, it will transform into itself or into a physically different state. A gauge transformation, by contrast, will only transform our description of a system from one gauge to another, without ever having any effect on the physical state of the system. Gauge transformations are comparable to rotations or translations of the coordinate system we use to describe a system. Another way of seeing the difference is by noting that it is possible to rotate or translate a superconductor in the laboratory, but as a matter of principle not possible to gauge transform it. Returning to the superconductor, I the student conclude that if the many different ground states only differ by a gauge transformation, they cannot be physically different. The ground state of a superconductor must hence be physically unique.

In fact, there is another way of looking at the problem which appears to confirm this conclusion. A BCS superconductor can not only be described in the grand-canonical ensemble, where the chemical potential rather than the number of particles is fixed, but also in the canonical ensemble, where the number of particles or pairs is fixed. Following Anderson [7], we can project out a (not normalized) state with N pairs from (1) via

$$|\psi_N\rangle = \int_0^{2\pi} d\phi \, e^{-iN\phi} |\psi_\phi\rangle \tag{3}$$

and obtain (see Appendix A)

$$|\psi_N\rangle = \int d^3\boldsymbol{x}_1 \cdots d^3\boldsymbol{x}_{2N} \varphi(\boldsymbol{x}_1 - \boldsymbol{x}_2) \cdots \varphi(\boldsymbol{x}_{2N-1} - \boldsymbol{x}_{2N})$$
$$\cdot \psi^\dagger_\uparrow(\boldsymbol{x}_1)\psi^\dagger_\downarrow(\boldsymbol{x}_2) \cdots \psi^\dagger_\uparrow(\boldsymbol{x}_{2N-1})\psi^\dagger_\downarrow(\boldsymbol{x}_{2N})|0\rangle, \tag{4}$$

where the real-space creation operator fields $\psi^\dagger_\sigma(\boldsymbol{x})$ are simply the Fourier transforms of the momentum-space creation operators $c^\dagger_{\boldsymbol{k}\sigma}$,

$$\psi^\dagger_\sigma(\boldsymbol{x}) = \frac{1}{\sqrt{V}} \sum_{\boldsymbol{k}} e^{-i\boldsymbol{k}\boldsymbol{x}} c^\dagger_{\boldsymbol{k}\sigma}, \quad c^\dagger_{\boldsymbol{k}\sigma} = \frac{1}{\sqrt{V}} \int d^3\boldsymbol{x} e^{i\boldsymbol{k}\boldsymbol{x}} \psi^\dagger_\sigma(\boldsymbol{x}). \tag{5}$$

The wave function for each of the individual pairs, which only depends on the relative coordinate, is (up to a normalization) given by

$$\varphi(\boldsymbol{x}) = \frac{1}{V} \sum_{\boldsymbol{k}} \frac{v_{\boldsymbol{k}}}{u_{\boldsymbol{k}}} e^{i\boldsymbol{k}\boldsymbol{x}}. \tag{6}$$

This form nicely illustrates that all the pairs have condensed into the same state, which is the essence of superfluidity. As $\varphi(\boldsymbol{x})$ is uniquely determined for a given Hamiltonian, the ground state (4) of a superconductor once more appears to be unique and non-degenerate.

So far the students train of thought. The conclusion reached is of course completely wrong: a superfluid, and in particular a superconductor, is characterized by a spontaneously broken symmetry, and, at least in the thermodynamic limit, there are many degenerate ground states. There are several mistakes in the students analysis. The first is his literal interpretation of the statement quoted in Section 1. In fact, *a gauge symmetry cannot spontaneously break down as a matter of principle*, *since it is not a physical symmetry of the system to begin with*, *but merely an invariance of description* [8]. The only way to violate a gauge symmetry is by choosing a gauge, which again has only an effect on our description, but not on the physical system itself.

In particular, the BCS ground state does not violate gauge invariance, even though statements to the contrary have been made in the literature. The apparent contradiction with (1) and (2) can be resolved by recalling that a gauge transformation only affects our description of the system, and is analogous to a rotation of the coordinate system we use in the example of a ferromagnet: if we rotate the coordinate system accordingly, a ground state with the magnetization vector pointing in the z-direction in the original coordinate system will "transform" into a state with the magnetization vector pointing in the x-direction in the new coordinate system, while the physical state has not been affected at all. So while the BCS wave function may look different in a different gauge, the state itself will remain the same.

It is worthwhile to rephrase this statement in equations. To begin with, let us consider a (relativistic quantum) field theory. Electromagnetic gauge invariance is the invariance of a given theory under $U(1)$ rotations of the complex scalar fields which carry the charge:

$$\psi^{\dagger}(x) \to e^{i\frac{e}{\hbar c}\Lambda(x)}\psi^{\dagger}(x), \quad \psi(x) \to e^{-i\frac{e}{\hbar c}\Lambda(x)}\psi(x), \tag{7}$$

where x denotes spacetime. We use the conventions $(x^{\mu}) = (x^0, x^1, x^2, x^3) = (ct, x, y, z)$, $x_{\mu} = g_{\mu\nu}x^{\nu}$, $1 = g_{00} = -g_{11} = -g_{22} = -g_{33}$. If the theory contains gradient terms in these fields (as it usually does), gauge invariance demands that they are minimally coupled to a $U(1)$ gauge field, i.e., the gradient terms must enter the Lagrangian as

$$\left(\partial_{\mu} + i\frac{e}{\hbar c}A_{\mu}(x)\right)\psi^{\dagger}(x) \quad \text{or} \quad \left(\partial_{\mu} - i\frac{e}{\hbar c}A_{\mu}(x)\right)\psi(x),$$

where $(\partial_{\mu}) \equiv (\partial/\partial x^{\mu}) = (\frac{1}{c}\partial_t, \nabla)$. The gauge field $(A_{\mu}) = (\Phi, -\boldsymbol{A})$ must transform according to

$$A_{\mu}(x) \to A_{\mu}(x) - \partial_{\mu}\Lambda(x). \tag{8}$$

The statement that the theory is gauge invariant simply means that the Lagrangian is invariant under the combined transformation (7) and (8). It is not a physical invariance, but an invariance of description, as it only amounts to a reparametrization of fields.

The concept of gauge invariance is implemented in a very similar way in non-relativistic quantum mechanics, where the gauge field $\boldsymbol{A}$ is no longer considered a dynamical variable, but an externally applied vector potential, and we usually do not describe a system by a Lagrange density, but by a Hamiltonian operator and its eigenstates. For pedagogical reasons, let us first assume a formulation in second quantization. Electromagnetic gauge invariance means once again that the description is invariant under $U(1)$ rotations of the particle creation and annihilation operator fields [9],

$$\psi^{\dagger}_{\sigma}(\boldsymbol{x}) \to e^{i\frac{e}{\hbar c}\Lambda(\boldsymbol{x})}\psi^{\dagger}_{\sigma}(\boldsymbol{x}), \quad \psi_{\sigma}(\boldsymbol{x}) \to e^{-i\frac{e}{\hbar c}\Lambda(\boldsymbol{x})}\psi_{\sigma}(\boldsymbol{x}). \tag{9}$$

The kinetic part of the Hamiltonian will again contain gradient terms in the operator fields, which once again must be minimally coupled to the electromagnetic gauge field. For example, the standard kinetic Hamiltonian for a quadratic dispersion

$$H_{\text{kin}} = \frac{1}{2m}\sum_{\sigma}\int d^3\boldsymbol{x}\psi^{\dagger}_{\sigma}(\boldsymbol{x})\left(-i\hbar\nabla + i\frac{e}{\hbar}\boldsymbol{A}(\boldsymbol{x})\right)^2\psi_{\sigma}(\boldsymbol{x}) \tag{10}$$

is obviously invariant under (9) provided we transform the gauge field simultaneously according to

$$\boldsymbol{A}(\boldsymbol{x}) \to \boldsymbol{A}(\boldsymbol{x}) + \nabla\Lambda(\boldsymbol{x}). \tag{11}$$

Let us now turn to the gauge transformation properties of the eigenstates. Consider a general N electron eigenstate

$$|\varphi\rangle = \sum_{\sigma_1\cdots\sigma_N}\int d^3\boldsymbol{x}_1\cdots d^3\boldsymbol{x}_N\varphi(\boldsymbol{x}_1\cdots\boldsymbol{x}_N;\sigma_1\cdots\sigma_N)\cdot\psi^{\dagger}_{\sigma_1}(\boldsymbol{x}_1)\cdots\psi^{\dagger}_{\sigma_N}(\boldsymbol{x}_N)|0\rangle. \tag{12}$$

The state is invariant under (9) provided we transform the wave function according to

$$\varphi(\boldsymbol{x}_1\cdots\boldsymbol{x}_N;\sigma_1\cdots\sigma_N) \to \prod_{j=1}^{N}e^{-i\frac{e}{\hbar c}\Lambda(\boldsymbol{x}_j)}\varphi(\boldsymbol{x}_1\cdots\boldsymbol{x}_N;\sigma_1\cdots\sigma_N). \tag{13}$$

This already illustrates the statement phrased in words above: a gauge transformation leaves physical states invariant. This is just not obvious in every formulation. If we formulate a problem in non-relativistic quantum mechanics in first quantization, a gauge transformation will only amount to (11) and (13), as we do not even introduce the operator fields $\psi^\dagger$ and ψ. As $\boldsymbol{A}(\boldsymbol{x})$ implements an externally applied magnetic field, we must choose a gauge in order to obtain explicit expressions for the Hamiltonian and the eigenstates. The vector potential and the wave functions will have different functional forms in different gauges. The gauge rotations (9) of the particle creation and annihilation operators, by contrast, only amount to a local change of variables; we could write

$$\psi_\sigma^\dagger(\boldsymbol{x}) \to \psi_\sigma^{\dagger\prime}(\boldsymbol{x}) = \mathrm{e}^{\mathrm{i}\frac{e}{\hbar c}\Lambda(\boldsymbol{x})}\psi_\sigma^\dagger(\boldsymbol{x}),$$

$$\psi_\sigma(\boldsymbol{x}) \to \psi_\sigma'(\boldsymbol{x}) = \mathrm{e}^{-\mathrm{i}\frac{e}{\hbar c}\Lambda(\boldsymbol{x})}\psi_\sigma(\boldsymbol{x}),$$

and then simply omit the primes. This part of the gauge transformation is often omitted as a choice of convention.

In the case of a BCS superconductor, such a convention would be all but propitious, as it would suggest that the ground state is not gauge invariant. The apparent contradiction in the students train of thought is immediately resolved as one uses the full and correct prescription for a gauge transformation,

$$c_{\boldsymbol{k}\sigma}^\dagger \to \mathrm{e}^{\mathrm{i}\frac{e}{\hbar c}\Lambda} c_{\boldsymbol{k}\sigma}^\dagger, \quad \phi \to \phi - \frac{2e}{\hbar c}\Lambda, \tag{14}$$

where the transformation of the phase ϕ is the equivalent of (13) above. Then the BCS ground state

$$\prod_{\boldsymbol{k}} (u_{\boldsymbol{k}} + v_{\boldsymbol{k}}\mathrm{e}^{\mathrm{i}\phi} c_{\boldsymbol{k}\uparrow}^\dagger c_{-\boldsymbol{k}\downarrow}^\dagger)|0\rangle,$$

is evidently gauge invariant; it is merely the label ϕ in $|\psi_\phi\rangle$ which will be adjusted under a gauge transformation. The transformation $\phi \to \phi - \frac{2e}{\hbar c}\Lambda$ is also required for the classical (or Ginzburg–Landau) order parameter field $\Psi^*(\boldsymbol{x})$, which is given by the expectation value of the operator field

$$\hat{\Psi}^\dagger(\boldsymbol{x}) \equiv \psi_\uparrow^\dagger(\boldsymbol{x})\psi_\downarrow^\dagger(\boldsymbol{x}), \tag{15}$$

to have the correct gauge transformation properties. The order parameter for the BCS ground state,

$$\Psi^*(\boldsymbol{x}) = \langle\psi_\phi|\psi_\uparrow^\dagger(\boldsymbol{x})\psi_\downarrow^\dagger(\boldsymbol{x})|\psi_\phi\rangle = \frac{1}{V}\sum_{\boldsymbol{k}} v_{\boldsymbol{k}}^* u_{\boldsymbol{k}}^* \mathrm{e}^{-\mathrm{i}\phi}, \tag{16}$$

transforms as a field of charge $-2e$ under (14):

$$\Psi^*(\boldsymbol{x}) \to \mathrm{e}^{\mathrm{i}\frac{2e}{\hbar c}\Lambda(\boldsymbol{x})}\Psi^*(\boldsymbol{x}). \tag{17}$$

This is the physically correct prescription. When we couple $\Psi^*(\boldsymbol{x})$ minimally to the electromagnetic gauge field, as required by (17), we obtain the correct effective field theory description of superconductivity. This theory displays the Higgs mechanism and yields London's equation. (By contrast, if we were to adhere to (2), $\Psi^*(\boldsymbol{x})$ would be invariant, could not be coupled to the electromagnetic gauge field, and no sensible effective field theory could be formulated.)

3. Order parameter considerations

Before proceeding further with the Higgs mechanism, I would like to return to the students train of thought and explain what is wrong with the conclusion he draws from the BCS wave function in position space. The problem here is that while the ground state is indeed unique for a finite system, there are many degenerate states in the thermodynamic limit, which correspond to different numbers of particles. To understand this issue in depth, it is best to first recall how rotational symmetry is spontaneously violated in ferromagnets and antiferromagnets. As a minimal model, we consider a three-dimensional cubic lattice of spins with spin quantum number S and assume the Heisenberg Hamiltonian [10]

$$H_J = J \sum_{\langle i,j \rangle} \boldsymbol{S}_i \boldsymbol{S}_j, \tag{18}$$

where the sum extends over all nearest-neighbor bonds $\langle i,j \rangle$ and $J < 0$ ($J > 0$) for a ferromagnet (an antiferromagnet).

In the case of a ferromagnet, the order parameter is given by the total spin operator

$$\boldsymbol{S}_{\text{tot}} \equiv \sum_i \boldsymbol{S}_i, \tag{19}$$

where the sum extends over all lattice sites. It commutes with the Hamiltonian,

$$[H_J, \boldsymbol{S}_{\text{tot}}] = 0, \tag{20}$$

and it is hence possible to choose simultaneous eigenstates of the Hamiltonian and the order parameter. In other words, the degenerate eigenstates of the order parameter corresponding to all possible directions the magnetization vector

$$\boldsymbol{M} = \langle \boldsymbol{S}_{\text{tot}} \rangle$$

can point to, are simultaneously degenerate eigenstates of the Hamiltonian. (For a ferromagnet with N spins, all the spins align and the ground states are just the states with maximal total spin $S_{\text{tot}} = NS$.) The ground state of the Hamiltonian is vastly (i.e., $2\,S_{\text{tot}} + 1$ fold) degenerate even if the system is finite.

The situation is different in the case of the antiferromagnet. The order parameter is given by the Néel vector, which in operator form is given by

$$\hat{\boldsymbol{N}} \equiv \sum_{i \in \text{A}} \boldsymbol{S}_i - \sum_{j \in \text{B}} \boldsymbol{S}_j, \tag{21}$$

where A and B denote the two sublattices of the (bipartite) cubic lattice. It does *not* commute with the Hamiltonian:

$$[H_J, \hat{\boldsymbol{N}}] \neq 0.$$

This implies that we cannot choose simultaneous eigenstates for the Hamiltonian and the order parameter. In fact, a theorem due to Marshall [11] states that the ground state for N even is unique and a spin singlet, or in other words, rotationally invariant. (It is possible to choose simultaneous eigenstates of H_J and $\boldsymbol{S}_{\text{tot}}$ as (20) holds independently of the sign of J.) The classical Néel order parameter,

$$\boldsymbol{N} = \langle \hat{\boldsymbol{N}} \rangle$$

will vanish for any finite system. This is not to say that there is no order for a finite system; it just manifests itself only through long-range correlations in the staggered spin–spin correlation function:

$$\langle \boldsymbol{S}_i \boldsymbol{S}_j \rangle \to \pm\text{const.} \quad \text{as } i - j \to \infty,$$

where the + sign applies for sites i and j on the same sublattice, the − sign for i and j on different sublattices, and $i - j \to \infty$ is understood to denote a very large separation within (the finite volume of) the system. As we approach the thermodynamic limit, the difference in energy between the lowest singlet and lowest eigenstates for $S_{\text{tot}} = 1,2,3,\ldots$ vanishes, and the ground state becomes degenerate (see Fig. 1A). These degenerate states can now be classified by the directions of the Néel vector $\boldsymbol{N}$, and the spontaneous breakdown of rotation symmetry is evident.

The situation in superconductors is analogous to the antiferromagnet: The (operator valued) order parameter (15) does not commute with the BCS Hamiltonian for any finite system even if we work in the grand-canonical ensemble, and the ground state for any finite volume will have a well defined particle number. The difference in energy between a system with N or $N \pm 1$ or $N \pm 2$, etc. pairs, however, will vanish in the thermodynamic limit (see Fig. 1B), and the many degenerate ground states can be classified by the phase ϕ of the (classical) order parameter

$$\Psi^*(\boldsymbol{x}) = \langle \hat{\Psi}^\dagger(\boldsymbol{x}) \rangle = |\Psi^*(\boldsymbol{x})| \mathrm{e}^{-\mathrm{i}\phi(\boldsymbol{x})}. \tag{22}$$

The broken symmetry is of course also present in a system with a fixed number of particles, but like in the case of the antiferromagnet, only as a long-range correlation of the (operator valued) order parameter field:

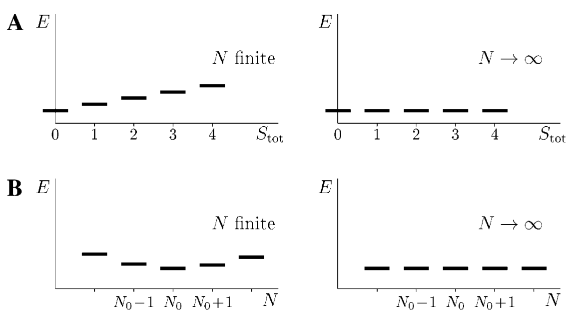


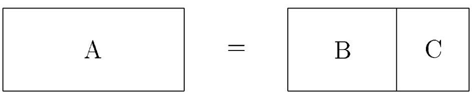
Fig. 1. In antiferromagnets (A) and superconductors (B), the ground state is unique for finite systems but degenerate in the thermodynamic limit.

$$\langle \hat{\Psi}^{\dagger}(\boldsymbol{x})\hat{\Psi}(\boldsymbol{y})\rangle \to \text{const.} \quad \text{as } |\boldsymbol{x}-\boldsymbol{y}| \to \infty, \tag{23}$$

where $\hat{\Psi}(\boldsymbol{x}) \equiv \psi_{\downarrow}(\boldsymbol{x})\psi_{\uparrow}(\boldsymbol{x})$ is simply the hermitian conjugate of $\hat{\Psi}^{\dagger}(\boldsymbol{x})$. This correlation is referred to as off diagonal long-range order (ODLRO) [12]. This type of order is characteristic to all superfluids, whether charged (like a superconductor) or neutral (like liquid helium), whether fermionic (like a superconductor or ^{3}He) or bosonic (like ^{4}He). For a bosonic superfluid, the (operator valued) order parameter $\hat{\Psi}^{\dagger}(\boldsymbol{x})$ and its hermitian conjugate $\hat{\Psi}(\boldsymbol{x})$ no longer create or annihilate a pair of fermions, but simply create or annihilate a single boson (like a ^{4}He atom).

The ODLRO is already evident from the position space wave function (4): Since all the pairs have condensed into the same quantum state, which is translationally invariant as it does not depend on the center-of-mass coordinates of the pairs, we expect to obtain a finite overlap with the original ground state if we rather clumsily (i.e., via $\hat{\Psi}(\boldsymbol{y})$) remove a pair of particles at some location $\boldsymbol{y}$ and equally clumsily (i.e., via $\hat{\Psi}^{\dagger}(\boldsymbol{x})$) recreate it at a distant location $\boldsymbol{x}$. In a superfluid or superconductor with a fixed number of particles, the phase ϕ will align over the entire system, like the direction of the staggered magnetization or Néel vector will align in an antiferromagnet.

To illustrate the significance of the phase once more, let us consider a large (but finite) superconductor A, and describe it as a combination of two superconductors B and C:



If we label the ground states of each superconductor by its phase, we can obviously write

$$|\psi_{\phi}^{\mathrm{A}}\rangle = |\psi_{\phi}^{\mathrm{B}}\rangle \otimes |\psi_{\phi}^{\mathrm{C}}\rangle$$

as the phase ϕ of the order parameter will align over the entire system. If we now transform to a description in terms of fixed numbers of pairs N_a for each superconductor,

$$|\psi_{N}^{a}\rangle = \int_{0}^{2\pi} \mathrm{d}\phi \mathrm{e}^{-\mathrm{i}N_a\phi}|\psi_{\phi}^{a}\rangle,$$

where a can be A, B, or C, the ground state of A is no longer a direct product of the ground states of B and C:

$$|\psi_{N_{\mathrm{A}}}^{\mathrm{A}}\rangle \neq |\psi_{N_{\mathrm{B}}}^{\mathrm{B}}\rangle \otimes |\psi_{N_{\mathrm{A}}-N_{\mathrm{B}}}^{\mathrm{C}}\rangle,$$

no matter how we choose N_{B}, as the phases no longer align. So while it is possible to describe a superconductor in a canonical ensemble (i.e., with fixed particle number), it is highly awkward to do so. It is comparable to a description of an antiferromagnet with long-range Néel order in terms of an overall singlet ground state of the system.

The most significant difference between an antiferromagnet and a superfluid or superconductor with regard to the order parameter is that the broken rotational symmetry in the former case is much more evident to us, as all the macroscopic objects in our daily life experience violate rotational symmetry at one level or another. In particular, the structure of the material in which antiferromagnetic order occurs provides us already with a reference frame for the direction the Néel order parameter may point to. In the case of the superconductor, we need a second superconductor to have a reference direction for the phase, and an interaction between the order parameter in both superconductors to detect a relative difference in the phases. (In practice, such an interaction may be accomplished by a pair tunneling or so-called Josephson junction.) The interference experiments will of course only be sensitive to the relative phase, and not the absolute phase in any of the superconductors, as all phases can, as a matter of principle, only be specified relative to some reference phase. In principle, the same is true for rotational invariance, but in this case the fixed stars provide us with a reference frame we perceive as "absolute."

We may conclude at this point that in a superfluid or superconductor, a symmetry is spontaneously violated, but this symmetry is *not* gauge invariance, but global $U(1)$ phase rotation symmetry. This is already evident from the fact that the discussion above made no reference to whether the order parameter field $\hat{\Psi}^{\dagger}(\boldsymbol{x})$ is charged or not, and equally well applies to neutral superfluids, where $\hat{\Psi}^{\dagger}(\boldsymbol{x})$ carries no charge.

There is, however, a very important difference between these two cases. If the order parameter field is neutral, the excitation spectrum of the system contains a gapless (or in the language of particle physics "massless") mode, a so-called Goldstone boson [1], which physically corresponds to very slow spatial variations in the direction (as for the case of broken rotational invariance) or phase (as for the case of a superfluid) of the classical order parameter field. If the order parameter field is charged, however, it couples to the electromagnetic gauge field, and the Goldstone boson is absent due to the Higgs mechanism. The physical principle underlying this mechanism was discovered by Anderson [13] in the context of superconductivity: as the electromagnetic interaction is long-ranged, the mode corresponding to very slow spatial variations in the phase ϕ of the superconducting order parameter, which implies currents by the equation of motion and hence also variations in the density of the superfluid by the continuity equation, acquires a gap (or "mass") given by the plasma frequency.

4. Effective field theory

Most of the phenomenology of superfluidity or superconductivity can be derived from a simple effective field theory, which in the latter case displays the Higgs mechanism. It is probably best to turn directly to the low-energy effective Lagrangian for the superconductor, as it contains the superfluid as the special case where the coupling e^{*} of the order parameter to the electromagnetic gauge field is set to zero. To motivate the Lagrangian, recall first the Ginzburg–Landau [14] expansion of the free energy density in terms of the order parameter (which is now normalized differently from (16) above) in the vicinity of the critical temperature T_c, where the transition between normal and superconducting phases occurs:

$$f(T,\Psi) = \frac{1}{2m^*}\left|\left(-i\hbar\nabla + \frac{e^*}{c}\boldsymbol{A}(\boldsymbol{x})\right)\Psi(\boldsymbol{x})\right|^2 + a(T)|\Psi(\boldsymbol{x})|^2 + \frac{1}{2}b(T)|\Psi(\boldsymbol{x})|^4 + \frac{1}{8\pi}\boldsymbol{B}(\boldsymbol{x})^2, \tag{24}$$

where m^* and $-e^* = -2e$ are the effective mass and charge of the electron pairs, respectively, and $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ is the magnetic field. The material parameter $a(T)$ changes sign to become negative as we pass through the transition from above, while $b(T)$ has to remain positive. Minimizing the free energy in the superconducting phase yields that: (i) the gradient term must vanish, (ii) $|\Psi(\boldsymbol{x})|^2 = -a/b$, and (iii) $\boldsymbol{B} = 0$. This means that the amplitude Ψ_0 of the order parameter $\Psi(\boldsymbol{x}) = \Psi_0 e^{i\phi}$ has to be fixed while the phase ϕ, which labels the many degenerate ground state configurations, can be arbitrary as long as the variation over the sample is given by

$$\nabla\phi = -\frac{e^*}{\hbar c}\boldsymbol{A},$$

which implies $\phi(\boldsymbol{x}) = \text{const.}$ if we choose the gauge $\boldsymbol{A}(\boldsymbol{x}) = 0$. In the vicinity of the transition, we may treat Ψ as a small parameter, which implies that the expansion (24) provides us with a complete description of the system at the level of thermodynamics.

The Ginzburg–Landau expansion is also helpful in motivating the low energy effective Lagrange density at low temperatures. To begin with, we may assume that since the amplitude fluctuations are massive, they do not enter in the low energy description. Taking $|\Psi(\boldsymbol{x})|$ to be constant, the free energy density above reduces to a constant, an electromagnetic field contribution, and

$$f_{\text{mag}} = \frac{n_s}{2m^*}\left(\hbar\nabla\phi(\boldsymbol{x}) + \frac{e^*}{c}\boldsymbol{A}(\boldsymbol{x})\right)^2, \tag{25}$$

where $n_s = |\Psi_0|^2$ is a phenomenological parameter which depends on the material and the temperature. It has the dimension of a density and is equal to the density of the superfluid in the absence of currents and inhomogeneities at $T = 0$, as we shall see below. It is usually referred to as the superfluid density, but it would be more appropriate to use the superfluid stiffness n_s/m^* as a parameter instead [7]. We will also see below that Galilean invariance of the superfluid implies that m^* is the bare mass of the superfluid particles, i.e., $m^* = 2m_e$ for Cooper pairs [15].

We take (25) to be part of the potential energy in the effective Lagrange density for the superfluid. The remaining contribution arises from the coupling of the charge of the superfluid to the electrostatic potential $\Phi(\boldsymbol{x})$, which is in leading order given by

$$f_{\text{el}} = -n_s e^* \Phi(\boldsymbol{x}).$$

This term is usually not included in the free energy of the superconductor, as it is always canceled off by another such term with opposite sign arising from the uniform positive background charge. It is essential to our effective field theory here, however, as it is part of the Lagrange density for the superfluid, while the uniform background charge is accounted for by another Lagrange density

$$\mathscr{L}_{\mathrm{b}}(x) = -n_{\mathrm{s}} e^* \Phi(x), \tag{26}$$

where $x = (ct, \boldsymbol{x})$ denotes spacetime. Note that f_{el} is not invariant under (time dependent) gauge transformations

$$\begin{aligned} \phi(x) &\to \phi(x) - \frac{e^*}{\hbar c} \Lambda(x), \\ \Phi(x) &\to \Phi(x) - \frac{1}{c} \partial_t \Lambda(x), \\ \boldsymbol{A}(x) &\to \boldsymbol{A}(x) + \nabla \Lambda(x). \end{aligned} \tag{27}$$

We now turn to the kinetic energy term in the effective Lagrange density. The simplest gauge invariant Lagrange density containing both potential energy terms above is given by [16]

$$\mathscr{L}_{\mathrm{s}}(x) = -n_{\mathrm{s}}(\hbar \partial_t \phi(x) - e^* \Phi(x)) - \frac{n_{\mathrm{s}}}{2m^*} \left(\hbar \nabla \phi(x) + \frac{e^*}{c} \boldsymbol{A}(x) \right)^2. \tag{28}$$

This Lagrange density, however, cannot be complete. The only term containing a time derivative, $\partial_t \phi(x)$, appears as a total derivative (here time derivative of ϕ) and hence does not affect the Euler–Lagrange equations of motion. It is nonetheless of physical significance, as it both ensures gauge invariance and accounts for the leading contribution to the particle density, as we will see below.

To obtain a second order time derivative term, recall that the characteristic feature of a neutral (i.e., $e^* = 0$) superfluid is that the only excitation at low energies is a sound wave with a linear dispersion

$$\omega(\boldsymbol{k}) = v |\boldsymbol{k}|, \tag{29}$$

where $\boldsymbol{k}$ is the wave number and v is the velocity of sound in the fluid. As we wish the effective Lagrange density for the superfluid both to be gauge invariant and to yield (29) as an equation of motion for $e^* = 0$, we arrive at

$$\begin{aligned} \mathscr{L}_{\mathrm{s}}(x) &= -n_{\mathrm{s}}(\hbar \partial_t \phi(x) - e^* \Phi(x)) \\ &\quad + \frac{n_{\mathrm{s}}}{2m^*} \left\{ \frac{1}{v^2} (\hbar \partial_t \phi(x) - e^* \Phi(x))^2 - \left(\hbar \nabla \phi(x) + \frac{e^*}{c} \boldsymbol{A}(x) \right)^2 \right\}. \end{aligned} \tag{30}$$

With

$$D_\mu \phi \equiv \hbar \partial_\mu \phi - \frac{e^*}{c} A_\mu, \tag{31}$$

where $(\partial_\mu) = (\frac{1}{c}\partial_t, \nabla)$ and $(A_\mu) = (\Phi, -\boldsymbol{A})$, the Lagrange density may also be written

$$\mathscr{L}_{\mathrm{s}} = -c n_{\mathrm{s}} D_0 \phi + \frac{n_{\mathrm{s}}}{2m^*} \left\{ \frac{c^2}{v^2} (D_0 \phi)^2 - (D_i \phi)^2 \right\}, \tag{32}$$

where $i = 1,2,3$. The total Lagrangian of the system is given by

$$L = \int \mathrm{d}^3 \boldsymbol{x} \{ \mathscr{L}_{\mathrm{s}}(x) + \mathscr{L}_{\mathrm{b}}(x) + \mathscr{L}_{\mathrm{em}}(x) \}, \tag{33}$$

where

$$\mathscr{L}_{\rm em} = -\frac{1}{16\pi} F_{\mu\nu} F^{\mu\nu} \quad \text{with } F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu \tag{34}$$

denotes the standard Maxwell Lagrange density for electromagnetism.

The astonishing feature is now that this simple Lagrangian for the compact $U(1)$ field $\phi(x)$ (compact since the values ϕ and $\phi + 2\pi$ describe the same physical state and hence must be identified) coupled to the electromagnetic gauge field accounts for all the essential features of superfluidity or superconductivity. There are also important corrections to it, but we will discover them automatically as we proceed.

To understand the physical content of the Lagrangian, it is highly instructive to study its symmetries, in particular particle number conservation and invariance under translations in space and time. We wish our analysis to apply both to the case of a neutral and a charged superfluid. In the former case, the theory is no longer invariant under a local $U(1)$ gauge transformation (as the electromagnetic gauge transformation (7) reduces to the identity transformation for $e^* = 0$) but still invariant under a global $U(1)$ rotation

$$\Psi(x) \to e^{i\lambda} \Psi(x) \quad \text{or} \quad \phi(x) \to \phi(x) + \lambda, \tag{35}$$

where λ is independent of spacetime. Physically, this symmetry corresponds to particle (or Cooper pair) number conservation. According to Noether's theorem [17], if under a given transformation the Lagrange density only changes by a total derivative,

$$D\mathscr{L}(x) \equiv \left. \frac{d\mathscr{L}(x, \lambda)}{d\lambda} \right|_{\lambda=0} = \partial_\mu F^\mu(x),$$

there is a conserved current associated with this symmetry:

$$J^\mu(x) = \text{const.} \cdot \left\{ \frac{\delta L_s}{\delta(\partial_\mu \phi(x))} D\phi(x) - F^\mu(x) \right\}, \tag{36}$$

where

$$D\phi(x) \equiv \left. \frac{d\phi(x, \lambda)}{d\lambda} \right|_{\lambda=0}.$$

Current conservation means $\partial_\mu J^\mu = 0$. Since (35) yields $F^\mu(x) = 0$ and $D\phi(x) = 1$, the particle four-current $(J^\mu) = (c\rho, \boldsymbol{J})$ is given by

$$J^\mu(x) = -\frac{1}{\hbar} \frac{\delta L_s}{\delta(\partial_\mu \phi(x))} = -\frac{\delta L_s}{\delta(D_\mu \phi(x))}, \tag{37}$$

where we have chosen the normalization such that the electric current equals the charge times the particle current:

$$-c \frac{\delta L_s}{\delta(A_\mu(x))} = -e^* J^\mu(x).$$

The Lagrange density (30) yields for the particle density

$$\rho(x) = -\frac{1}{\hbar}\frac{\delta L_{\rm s}}{\delta(\partial_t\phi(x))} \tag{38}$$

$$= n_{\rm s} - \frac{n_{\rm s}}{m^*}\frac{1}{v^2}(\hbar\partial_t\phi(x) - e^*\Phi(x)) \tag{39}$$

and for the particle current

$$\boldsymbol{J}(x) = -\frac{1}{\hbar}\frac{\delta L_{\rm s}}{\delta(\nabla\phi(x))} \tag{40}$$

$$= \frac{n_{\rm s}}{m^*}\left(\hbar\nabla\phi(x) + \frac{e^*}{c}\boldsymbol{A}(x)\right). \tag{41}$$

The corresponding conservation law is just the continuity equation

$$\partial_t\rho + \nabla\boldsymbol{J} = 0. \tag{42}$$

Note that since our Lagrangian (33) does not depend on $\phi(x)$, but only on derivatives of $\phi(x)$, i.e.,

$$\frac{\delta L_{\rm s}}{\delta(\phi(x))} = 0,$$

Eq. (37) implies that the current conservation law

$$-\hbar\partial_\mu J^\mu(x) = \partial_\mu\frac{\delta L_{\rm s}}{\delta(\partial_\mu\phi(x))} = 0$$

is equivalent the Euler–Lagrange equation for the field $\phi(x)$. For a neutral superfluid, we obtain

$$\left(\frac{1}{v^2}\partial_t^2 - \nabla^2\right)\phi(x) = 0 \tag{43}$$

and hence the dispersion (29) by Fourier transformation.

The most important implication of (37) for the particle four-current is, however, that the density $\rho(x)$ is up to a numerical factor equal to the momentum field $\pi(x)$ conjugate to $\phi(x)$:

$$-\hbar\rho(x) = \pi(x) \equiv \frac{\delta L_{\rm s}}{\delta(\partial_t\phi(x))}. \tag{44}$$

We may hence go over to an Hamiltonian formulation, and write the Hamiltonian density

$$\mathscr{H}_{\rm s}(x) \equiv -\hbar\rho(x)\partial_t\phi(x) - \mathscr{L}_{\rm s}(x), \tag{45}$$

which is now considered a functional of $\rho(x)$, $\partial_i\phi(x)$, $\Phi(x)$, and $A_i(x)$, but not $\partial_t\phi(x)$. (In principle, $\mathscr{H}_{\rm s}$ could also depend through $\mathscr{L}_{\rm s}$ on $\phi(x)$ and x. Note also that (45) as the generator of time translations is not invariant under time dependent gauge transformations, while the equations of motions below are invariant.) The Hamiltonian is of course given by

$$H_{\mathrm{s}} = \int \mathrm{d}^3\boldsymbol{x}\mathscr{H}_{\mathrm{s}}(x). \tag{46}$$

Hamilton's equations are in analogy to the familiar equations

$$\dot{q} = \frac{\partial H(p,q)}{\partial p}, \quad \dot{p} = -\frac{\partial H(p,q)}{\partial q}$$

from classical mechanics given by

$$\partial_t \phi(x) = \frac{\delta H_{\mathrm{s}}}{\delta(\pi(x))} = -\frac{1}{\hbar}\frac{\delta H_{\mathrm{s}}}{\delta(\rho(x))} \tag{47}$$

and

$$-\hbar\partial_t\rho(x) = \partial_t\pi(x) = \partial_i\frac{\delta H_{\mathrm{s}}}{\delta(\partial_i\phi(x))} - \frac{\delta H_{\mathrm{s}}}{\delta(\phi(x))}. \tag{48}$$

With regard to the explicit equations of motion for the fields, these equations are equivalent to the Euler–Lagrange equation. They provide, however, additional information regarding the physical interpretation. To extract this information, it is propitious to study the other conservation laws corresponding to energy and momentum first.

The theory is invariant under spacetime translations $x \to x - e\lambda$, where e is an arbitrary unit vector in spacetime (e.g. $e = (1, 0, 0, 0)$ or $e^\nu = \delta_{0\nu}$ for a translation in time). The infinitesimal translations are equivalent to the field and density transformations [17]

$$\begin{aligned}
&\phi(x) \to \phi(x + e\lambda) = \phi(x) + \lambda e^\nu\partial_\nu\phi(x),\\
&A_\mu(x) \to A_\mu(x + e\lambda) = A_\mu(x) + \lambda e^\nu\partial_\nu A_\mu(x),\\
&\mathscr{L}(x) \to \mathscr{L}(x + e\lambda) = \mathscr{L}(x) + \lambda\partial_\nu(e^\nu\mathscr{L}(x)),
\end{aligned}$$

which implies $D\phi = e^\nu\partial_\nu\phi$, $DA_\mu = e^\nu\partial_\nu A_\mu$, and $F^\mu(x) = e^\mu\mathscr{L}(x)$. The conserved current associated with this symmetry is according to (36) given by

$$J^\mu = \frac{\delta L}{\delta(\partial_\mu\phi)}e^\nu\partial_\nu\phi + \frac{\delta L}{\delta(\partial_\mu A_\kappa)}e^\nu\partial_\nu A_\kappa - e^\mu\mathscr{L} = e_\nu T^{\mu\nu}_{\mathrm{can}},$$

where the canonical energy–momentum tensor $T^{\mu\nu}_{\mathrm{can}}$ is the sum of the contributions from the superfluid, the uniformly charged background, and the electromagnetic field:

$$T^{\mu\nu}_{\mathrm{can}} = T^{\mu\nu}_{\mathrm{s,can}} + T^{\mu\nu}_{\mathrm{b,can}} + T^{\mu\nu}_{\mathrm{em,can}},$$

where

$$T^{\mu\nu}_{\mathrm{s,can}} = \frac{\delta L_{\mathrm{s}}}{\delta(\partial_\mu\phi)}\partial^\nu\phi - g^{\mu\nu}\mathscr{L}_{\mathrm{s}}, \tag{49}$$

$$T^{\mu\nu}_{\mathrm{b,can}} = -g^{\mu\nu}\mathscr{L}_{\mathrm{b}},$$

$$T^{\mu\nu}_{\mathrm{em,can}} = \frac{\delta L_{\mathrm{em}}}{\delta(\partial_\mu A_\kappa)}\partial^\nu A_\kappa - g^{\mu\nu}\mathscr{L}_{\mathrm{em}}. \tag{50}$$

The conservation law $\partial_\mu T^{\mu\nu}_{\rm can} = 0$ describes energy conservation for $\nu = 0$ and momentum conservation for $\nu = i$. The $\mu = 0$ components of $T^{\mu\nu}_{\rm can}$ correspond to energy and momentum densities; in particular, $-c\int {\rm d}^3x T^{00}_{\rm can}$ generates translations in time and $\int {\rm d}^3x T^{0i}_{\rm can}$ translations in space.

In the case of a gauge theory, like the theory of a charged superfluid we consider here, it is not possible to interpret $T^{00}_{\rm s,can}$ as the energy or $\frac{1}{c}T^{0i}_{\rm s,can}$ as the kinematical momentum density of the superfluid. The reason is simply that (49) (and also (50)) is not gauge invariant. To circumvent this problem, we simply supplement the naive translations by a suitable gauge transformation, such that the fields transform covariantly:

$$\begin{aligned} \phi &\to \phi + \lambda e^\nu\left(\partial_\nu\phi - \frac{e^*}{\hbar c}A_\nu\right), \\ A_\kappa &\to A_\kappa + \lambda e^\nu(\partial_\nu A_\kappa - \partial_\kappa A_\nu). \end{aligned}$$

The gauge transformation is hence given by (27) with $\Lambda(x) = \lambda e^\nu A_\nu(x)$. This yields the "kinematical" energy momentum tensor, with contributions from the superfluid and the electromagnetic field

$$\begin{aligned} T^{\mu\nu}_{\rm s} &= \frac{\delta L_{\rm s}}{\delta(D_\mu\phi)}D^\nu\phi - g^{\mu\nu}\mathscr{L}_{\rm s}, \\ T^{\mu\nu}_{\rm em} &= \frac{\delta L_{\rm em}}{\delta(F_{\mu\kappa})}F^\nu_{\ \kappa} - g^{\mu\nu}\mathscr{L}_{\rm em}. \end{aligned} \tag{51}$$

These expressions are manifestly gauge invariant. Using (37) with $\mu = 0$ or (38), we can write the energy density of the superfluid

$$T^{00}_{\rm s}(x) = -\rho(x)(\hbar\partial_t\phi(x) - e^*\Phi(x)) - \mathscr{L}_{\rm s}(x). \tag{52}$$

Note that this is numerically equal to

$$T^{00}_{\rm s}(x) = \mathscr{H}_{\rm s}(x) + \rho(x)e^*\Phi(x). \tag{53}$$

Similarly, we can write the momentum density

$$\frac{1}{c}T^{0i}_{\rm s}(x) = \rho(x)\left(\hbar\nabla\phi(x) + \frac{e^*}{c}\boldsymbol{A}(x)\right). \tag{54}$$

We can use this expression to introduce the superfluid velocity $\boldsymbol{v}_{\rm s}(x)$. In terms of $\boldsymbol{v}_{\rm s}(x)$, the momentum density of the superfluid has to be given by

$$\frac{1}{c}T^{0i}_{\rm s}(x) \stackrel{!}{=} \rho(x)m^*\boldsymbol{v}_{\rm s}(x), \tag{55}$$

which leads us to define

$$m^*\boldsymbol{v}_{\rm s}(x) \equiv \hbar\nabla\phi(x) + \frac{e^*}{c}\boldsymbol{A}(x) = D_i\phi(x). \tag{56}$$

Since $\boldsymbol{v}_{\rm s}(x)$ is to be interpreted as a physical velocity, it has to transform like a velocity under a Galilean transformation,

$$\boldsymbol{v}_{\rm s}(x) \to \boldsymbol{v}_{\rm s}(x) + \boldsymbol{u}. \tag{57}$$

The total momentum of the superfluid will hence transform according to

$$\frac{1}{c}\int \mathrm{d}^3\boldsymbol{x} T_{\mathrm{s}}^{0i}(x) \to \frac{1}{c}\int \mathrm{d}^3\boldsymbol{x} T_{\mathrm{s}}^{0i}(x) + \int \mathrm{d}^3\boldsymbol{x}\rho(x)m^*\boldsymbol{u}, \tag{58}$$

which implies directly that in a translationally invariant system, m^* has to be the bare mass of the superfluid particles or Cooper pairs [18].

It should also be possible to express the particle current in terms of the superfluid velocity. Since the same particles which carry the momentum also carry the current, the particle current has to be given by

$$\boldsymbol{J}(x) = \rho(x)\boldsymbol{v}_{\mathrm{s}}(x). \tag{59}$$

This is almost, but not quite, equivalent to our earlier expression (41), as n_{s} is only the leading contribution to $\rho(x)$. So either (59) with (56) or (41) is not fully correct. To see which one, recall that we have only used the general expression for the density (38) as defined through particle number conservation in obtaining (54) and hence (56) and (59) from (51), while we have used the explicit expression for the Lagrange density (30) in obtaining (41). In other words, only symmetry considerations enter in (54), while (41) depends explicitly on the Lagrange density. The expression for the momentum density (54), and hence our definition of the superfluid velocity (56), is therefore exact, while the expression (41) for the particle current is only an approximation [19].

The expression for the current, however, will assume the exact and physically correct form (59) if we introduce suitable corrections to the effective Lagrangian. To obtain these, we simply require the Lagrangian to satisfy [20]

$$\frac{1}{c}T_{\mathrm{s}}^{0i}(x) = m^* J^i(x) \tag{60}$$

or

$$\frac{1}{c}\frac{\delta L_{\mathrm{s}}}{\delta(D_0\phi)}D^i\phi = -m^*\frac{\delta L_{\mathrm{s}}}{\delta(D_i\phi)}. \tag{61}$$

Upon integration of this equation we find that the Lagrange density must be of the form

$$\mathscr{L}_{\mathrm{s}} = P\left(cD_0\phi + \frac{1}{2m^*}(D_i\phi)^2\right), \tag{62}$$

where P is an arbitrary polynomial. Our superfluid Lagrange density (32) will assume this form if we add third and fourth order corrections in $D_\mu\phi$; the full superfluid Lagrange density is then given by [20,21]

$$\mathscr{L}_{\mathrm{s}} = -n_{\mathrm{s}}\left(cD_0\phi + \frac{1}{2m^*}(D_i\phi)^2\right) + \frac{n_{\mathrm{s}}}{2m^*}\frac{1}{v^2}\left(cD_0\phi + \frac{1}{2m^*}(D_i\phi)^2\right)^2. \tag{63}$$

It yields for the particle density

$$\rho(x) = -\frac{1}{c}\frac{\delta L_{\rm s}}{\delta(D_0\phi)} = n_{\rm s} - \frac{n_{\rm s}}{m^*}\frac{1}{v^2}\left(cD_0\phi + \frac{1}{2m^*}(D_i\phi)^2\right) = n_{\rm s} - \frac{n_{\rm s}}{m^*}\frac{1}{v^2}\left\{(\hbar\partial_t\phi - e^*\Phi) + \frac{1}{2m^*}\left(\hbar\nabla\phi + \frac{e^*}{c}\boldsymbol{A}\right)^2\right\} \tag{64}$$

and for the particle current

$$\boldsymbol{J}(x) = -\frac{\delta L_{\rm s}}{\delta(D_i\phi)} = \rho(x)\boldsymbol{v}_{\rm s}(x), \tag{65}$$

where $\rho(x)$ and $\boldsymbol{v}_{\rm s}(x)$ are given by (64) and (56).

Let us now return to Hamilton's equations, and in particular their physical interpretation. With (53) we may rewrite (47) as

$$\hbar\partial_t\phi(x) = -\frac{\delta H_{\rm s}}{\delta(\rho(x))} = -\frac{\partial\mathscr{H}_{\rm s}(x)}{\partial\rho(x)} = -\frac{\partial T_{\rm s}^{00}(x)}{\partial\rho(x)} + e^*\Phi(x) = -\mu(x) + e^*\Phi(x), \tag{66}$$

where we have used the definition of the chemical potential. This is one of two equations Anderson [7] refers to as "characteristic of superfluidity." In analogy to the definition (56) of $\boldsymbol{v}_{\rm s}(x)$, we rewrite it for later purposes as

$$-\mu(x) = \hbar\partial_t\phi(x) - e^*\Phi(x) = cD_0\phi(x). \tag{67}$$

Taking the gradient and adding $\frac{e^*}{c}\partial_t\boldsymbol{A}(x)$ on both sides of (66), we obtain

$$\partial_t\left(\hbar\nabla\phi(x) + \frac{e^*}{c}\boldsymbol{A}(x)\right) = -\nabla\mu(x) - e^*\boldsymbol{E}(x) = -\nabla\mu_{\rm el.chem.}(x), \tag{68}$$

where we used the definitions of the electric field,

$$\boldsymbol{E} \equiv -\nabla\Phi - \frac{1}{c}\partial_t\boldsymbol{A},$$

and of the electrochemical potential. With (56) we may write

$$m^*\partial_t\boldsymbol{v}_{\rm s}(x) = -\nabla\mu_{\rm el.chem.}(x). \tag{69}$$

The gradient of the electrochemical potential (or chemical potential for a neutral superfluid) is usually [7] identified with minus the total force on the particles, and (69) is referred to as the "acceleration equation." This is, however, not quite correct. $\partial_t\boldsymbol{v}_{\rm s}$ in (69) denotes the time derivative in the superfluid velocity field at spacetime x (known as "local acceleration" in hydrodynamics), while the force on the particles is given by the time derivative of the velocity of a given particle in the fluid at x ("substantial acceleration" in hydrodynamics):

$$\frac{1}{m^*}\boldsymbol{F} = \frac{{\rm d}\boldsymbol{v}_{\rm s}}{{\rm d}t} = \partial_t\boldsymbol{v}_{\rm s} + (\boldsymbol{v}_{\rm s}\nabla)\boldsymbol{v}_{\rm s}. \tag{70}$$

Nonetheless, (69) is one of the fundamental equations in the phenomenology of superfluidity. It states that if there is a gradient in the electrochemical (or chemical) potential in a superconductor (or superfluid), the superfluid will be "accelerated" without any frictional damping. On the other hand, if the superfluid flow is stationary, the electrochemical (or chemical) potential has to be constant across the superconductor (or superfluid). Since a voltmeter measures a difference in the electrochemical potential, there cannot be a voltage across a superconductor unless the flow is "accelerated."

Let us now turn to Hamilton's second equation (48). We first rewrite it as

$$-\hbar\partial_t\rho(x) = \partial_i \frac{\delta H_s}{\delta(\partial_i\phi(x))} - \frac{\partial\mathscr{H}_s(x)}{\partial\phi(x)}. \tag{71}$$

Since the last term in (53), $\rho(x)\, e^*\Phi(x)$, does not depend on the phase $\phi(x)$, we may replace $\mathscr{H}_s(x)$ in the last term in (71) by $T_s^{00}(x)$. Integrating the resulting equation over the superfluid, discarding a boundary term, and defining a "global" derivative with respect to the phase,

$$\frac{\partial F[\phi(x)]}{\partial\phi} \equiv \lim_{\Delta\phi\to 0} \frac{F[\phi(x)+\Delta\phi]}{\Delta\phi},$$

where $F[\phi(x)]$ is an arbitrary functional of $\phi(x)$ and $\Delta\phi$ an infinitesimal independent of spacetime, yields

$$-\hbar\partial_t \int \mathrm{d}^3\boldsymbol{x}\rho(x) = -\frac{\partial}{\partial\phi}\int \mathrm{d}^3\boldsymbol{x} T_s^{00}(x)$$

or

$$\hbar\partial_t N = \frac{\partial E_s}{\partial\phi}, \tag{72}$$

where N is the number of particles (or pairs) and E_s the energy of the superfluid. This is the other "characteristic equation" of superfluidity [7].

This concludes our derivation or motivation of the fundamental equations of superfluidity in the limit of low temperatures and low energies. Before turning to the phenomenology these equations imply, I would like to digress briefly and justify one of the implicit assumptions made above. The assumption is that we can describe the macroscopic quantum phenomena of superfluidity with a classical effective field theory, or in other words, that we may consider both the phase $\phi(x)$ and its conjugate field $\pi(x)=-\hbar\rho(x)$ as thermodynamic variables. To justify this assumption, let us canonically quantize the theory by imposing

$$[\hat{\phi}(\boldsymbol{x},t), \hat{\pi}(\boldsymbol{y},t)] = \mathrm{i}\hbar\delta(\boldsymbol{x}-\boldsymbol{y}). \tag{73}$$

Integration of $\boldsymbol{y}$ over the superfluid yields

$$[\hat{\phi}(\boldsymbol{x}), \hat{N}] = -\mathrm{i},$$

which in turn implies the uncertainty relation

$$\Delta\phi(\boldsymbol{x})\Delta N \geqslant \tfrac{1}{2}.$$

If we assume that the number of particles in the superfluid takes on a macroscopic value of order $N \approx 10^{20}$, a ΔN of the order of $\sqrt{N}$ implies a relative uncertainty of order 10^{-10} in the particle number and the phase. These numbers are comparable to the position and momentum uncertainties of a macroscopic object. The description of a macroscopic superfluid in terms of a classical field theory is therefore as appropriate as the classical description of any other macroscopic object. This is of course not in contradiction with the fact that Planck's constant $\hbar$ appears in this effective field theory. We will see below that it manifests itself in a family of "quantum effects," which are related to the compactness of the $U(1)$ field $\phi(x)$. These effects require either a non-trivial topology or more than one superfluid, and are very similar for neutral and for charged superfluids.

5. Phenomenology and the Higgs mechanism

To begin with, however, let us consider the superfluid flow in a simply connected superfluid. The phenomenology depends strikingly on whether the fluid is charged or not. For a neutral superfluid, $e^* = 0$, and the gauge field decouples completely. Even for a fixed set of boundary conditions, we have an infinite set of solutions for the superfluid flow, corresponding via

$$m^* \boldsymbol{v}_\mathrm{s}(x) = \hbar \nabla \phi(x)$$

to all possible choices of the phase field $\phi(x)$. In a simply connected superfluid, the flow will be vortex-free, i.e.,

$$\nabla \times \boldsymbol{v}_\mathrm{s}(x) = 0,$$

and subject to boundary constraints, but apart from this, it only has to satisfy the continuity equation as an equation of motion.

The simplest example of a multiply connected superfluid is a superfluid with a line defect, or vortex, along which the magnitude $|\Psi(x)|$ of the superfluid order parameter vanishes. The phase $\phi(x)$ still has to be single valued everywhere in the fluid, but being a phase, its value may change by a multiple of 2π as we circumvent the line defect along a closed curve ∂S:

$$\oint_{\partial S} \nabla \phi(x)\, \mathrm{d}\boldsymbol{l} = 2\pi n, \tag{74}$$

where n is an integer. The angular momentum of each superfluid "particle" around the vortex is hence quantized in units of $\hbar$. With Stokes theorem and the definition

$$\boldsymbol{\omega}(x) = \nabla \times \boldsymbol{v}_\mathrm{s}(x), \tag{75}$$

we may express this alternatively as quantization condition for the vorticity

$$\int_S \boldsymbol{\omega}(x) \cdot \boldsymbol{n}\, \mathrm{d}a = \frac{2\pi\hbar n}{2m^*}, \tag{76}$$

where $\boldsymbol{n}$ is a unit vector normal to the surface and the area integral extends over any open surface S which is pierced by the vortex once. The quantization of vortices in a superfluid is the simplest of the "quantum effects" alluded to above, where Planck's constant $\hbar$ enters in the phenomenology through the compactness of the field $\phi(x)$. (If $\phi(x)$ was not compact, we could eliminate $\hbar$ completely from the effective theory by rescaling $\phi(x) \to \hbar\phi(x)$.)

Let us now turn to the phenomenology of a simply connected charged superfluid or superconductor, which displays the Higgs mechanism. The essence of the mechanism is that the phase field $\phi(x)$ looses its independent significance in the presence of the gauge field. There are two ways of seeing this. The first is on the level of the equations of motion. We can simply choose a gauge such that $\phi(x) = 0$ everywhere in the fluid; any other choice of gauge can be brought into this gauge via (27) with

$$\Lambda(x) = \frac{\hbar c}{e^*}\phi(x).$$

The second is on the level of the effective Lagrangian. We may introduce a new vector field

$$-\frac{e^*}{c}A'_\mu \equiv D_\mu\phi = \hbar\partial_\mu\phi - \frac{e^*}{c}A_\mu. \tag{77}$$

In terms of this field, the superfluid Lagrange density (63) looks the same except that all terms containing ϕ have disappeared. In particular, the terms quadratic in the derivatives of ϕ in $\mathscr{L}_\mathrm{s}$ have turned into a mass term

$$\frac{n_\mathrm{s}}{2m^*}\left\{\frac{1}{v^2}\left(e^*A'_0\right)^2 - \left(\frac{e^*}{c}\boldsymbol{A}'\right)^2\right\} \tag{78}$$

for the vector field. The Maxwell Lagrange density and the Lagrange density for the uniform neutralizing background charge take the same form with $F_{\mu\nu}$ and A_μ replaced by $F'_{\mu\nu}$ and A'_μ, respectively, except for a total derivative or boundary term we discard. Thus the massless gauge field A_μ is replaced by a massive vector field A'_μ, while the Goldstone boson ϕ disappeared. The total number of degrees of freedom, however, is preserved: before, the massless vector field has two (the two helicity states of the photon) and the Goldstone boson one, while the massive vector field after the change of variables has three degrees of freedom. In Sidney Coleman's words, "the vector field has eaten the Goldstone bosons and grown heavy" [1]. We will return to this issue after studying the phenomenology of the superconductor using the equations of motion.

The Euler–Lagrange equation for A_μ,

$$\partial_\mu\frac{\delta L}{\delta(\partial_\mu A_\nu)} - \frac{\delta L}{\delta A_\nu} = 0,$$

yields Maxwell's electrodynamics with electric charge density $-e^*(\rho - n_\mathrm{s})$ and current $-e^*\mathbf{J}$,

$$\nabla\cdot\boldsymbol{E} = -4\pi e^*(\rho - n_\mathrm{s}), \tag{79}$$

$$\nabla\times\boldsymbol{B} - \frac{1}{c}\partial_t\boldsymbol{E} = -\frac{4\pi e^*}{c}\boldsymbol{J}, \tag{80}$$

where

$$\boldsymbol{E} = -\nabla\Phi - \frac{1}{c}\partial_t \boldsymbol{A},\ \ \boldsymbol{B} = \nabla \times \boldsymbol{A},$$

and ρ and $\boldsymbol{J}$ are given by (64) and (65), respectively. In principle, we could also obtain the continuity equation (42) as the Euler–Lagrange equation for ϕ, but since L depends on derivatives of A_μ only through $F_{\mu\nu}$ and ϕ is minimally coupled to A_μ, (42) is automatically satisfied by any solution of (79) and (80). This is consistent with the fact that ϕ has lost its independent significance due to the Higgs mechanism.

For convenience, we choose the gauge $\phi(x) = 0$. Then (64), (65), and (56) imply

$$4\pi e^*(\rho - n_s) = \frac{4\pi e^{*2} n_s}{m^* v^2}\left(\Phi - \frac{e^*}{2m^*c^2}\boldsymbol{A}^2\right), \tag{81}$$

$$\frac{4\pi e^*}{c}\boldsymbol{J} = \frac{4\pi e^{*2} n_s}{m^* c^2}\boldsymbol{A}\left\{1 + \frac{e^*}{m^* v^2}\left(\Phi - \frac{e^*}{2m^*c^2}\boldsymbol{A}^2\right)\right\}. \tag{82}$$

Let us now restrict our attention to quasistatic phenomena, where we can neglect the time derivative terms. The analysis given below implies that this assumption holds for frequencies significantly smaller than c/λ_L, where

$$\lambda_L = \sqrt{\frac{m^* c^2}{4\pi e^{*2} n_s}} \tag{83}$$

is the London penetration depth. Then (79)–(82) reduce to

$$\nabla^2\Phi = \frac{c^2}{v^2}\frac{1}{\lambda_L^2}\left(\Phi - \frac{e^*}{2m^*c^2}\boldsymbol{A}^2\right), \tag{84}$$

$$\nabla^2\boldsymbol{A} - \nabla(\nabla\boldsymbol{A}) = \frac{1}{\lambda_L^2}\boldsymbol{A}\left\{1 + \frac{e^*}{m^* v^2}\left(\Phi - \frac{e^*}{2m^*c^2}\boldsymbol{A}^2\right)\right\}. \tag{85}$$

Let us first look at the linear terms in these equations, i.e., the solution for infinitesimal Φ and $\boldsymbol{A}$. Under quasistatic conditions, (80) implies $\nabla\boldsymbol{J} = 0$ and with (82) for infinitesimal fields $\nabla\boldsymbol{A} = 0$. The equations reduce to

$$\nabla^2\Phi = \frac{c^2}{v^2}\frac{1}{\lambda_L^2}\Phi, \tag{86}$$

$$\nabla^2\boldsymbol{A} = \frac{1}{\lambda_L^2}\boldsymbol{A}, \tag{87}$$

i.e., we have electric screening in addition to magnetic screening, but with a screening length reduced by a factor v/c. This leads us to conjecture that the dominant energy is the Coulomb interaction, which effects charge neutrality or $\rho(x) \approx n_s$. We now simply assume that this is a valid approximation, and justify it a posteriori. Then (81) implies

$$\Phi - \frac{e^*}{2m^*c^2}\boldsymbol{A}^2 = 0, \tag{88}$$

and (82) reduces to

$$\boldsymbol{J} = \frac{e^* n_s}{m^* c} \boldsymbol{A}. \tag{89}$$

Taking the curl of this equation, we obtain London's equation [22,23]

$$\nabla \times \boldsymbol{J} = \frac{e^* n_s}{m^* c} \boldsymbol{B}. \tag{90}$$

Under quasistatic conditions, we have again $\nabla \boldsymbol{J} = 0$ and with (89) $\nabla \boldsymbol{A} = 0$, which implies that (85) reduces to (87). The solution of (87) describes exponential screening with penetration depth λ_L. If we have, for example, a superconductor which occupies the half-space $x > 0$ subject to an external magnetic field $\boldsymbol{B} = B_0 \hat{\boldsymbol{y}}$ at the boundary $x = 0$, we obtain

$$\boldsymbol{A} = A_0 \mathrm{e}^{-x/\lambda_L} \hat{\boldsymbol{z}}, \quad \boldsymbol{B} = B_0 \mathrm{e}^{-x/\lambda_L} \hat{\boldsymbol{y}}, \quad \boldsymbol{J} = J_0 \mathrm{e}^{-x/\lambda_L} \hat{\boldsymbol{z}}, \tag{91}$$

where

$$A_0 = \lambda_L B_0, \quad J_0 = \sqrt{\frac{n_s}{4\pi m^*}} B_0.$$

The screening of the magnetic field is known as the Meissner effect. According to (88), the vector potential implies an electrostatic potential

$$\Phi = \frac{B_0^2}{8\pi e^* n_s} \mathrm{e}^{-2x/\lambda_L}. \tag{92}$$

This potential allows us to verify the validity of our approximation $\rho(x) \approx n_s$. Substituting (92) into (84), we find that the ratio of the neglected term to the terms kept is

$$\frac{\nabla^2 \Phi}{\frac{c^2}{v^2} \frac{1}{\lambda_L^2} \Phi} = \frac{4v^2}{c^2} \ll 1, \tag{93}$$

i.e., the approximation is excellent.

The electrostatic potential (92) is called the London or Bernoulli Hall effect [22,20]. To understand its physical origin, it is best to rewrite (88) with (56) for $\phi(x) = 0$ in terms of the superfluid velocity:

$$-e^* \Phi + \frac{1}{2} m^* \boldsymbol{v}_s^2 = 0. \tag{94}$$

The electrostatic potential simply compensates the kinetic energy contribution to the chemical potential, as required by (66) with $\phi(x) = 0$. For stationary flow, this condition reduces to the requirement that the electrochemical potential $\mu_{\text{el.chem.}}$ is constant across the superconductor. In practice, the London Hall effect can only be measured with capacitive contacts, as ohmic contacts are sensitive to the electrochemical rather than the electrostatic potential [24]. The effect furnishes us with an independent meaning of the superfluid density n_s or the effective mass m^*, while under quasistatic conditions all other effects [25] depend only on the superfluid stiffness n_s/m^*. The underlying theoretical reason is that the London Hall effect is a consequence of the corrections incorporated in the effective Lagrange density (63). In these terms, the parameter m^* enters by itself, while (apart from a total time derivative term irrelevant to the equations of motion) only the combination n_s/m^* entered in the previous approximative Lagrangian (33) with (30), (26), and (34).

Since we have given a precise definition of the superfluid velocity, it is legitimate to ask whether the London Hall effect (94) balances the Lorentz force, or, if not, what other forces balance it. The total electromagnetic force on a given particle with charge $-e^*$ in the fluid is given by

$$\boldsymbol{F}_{\text{em}} = -e^*\left(\boldsymbol{E} + \frac{\boldsymbol{v}_s}{c} \times \boldsymbol{B}\right) = -e^*\left(-\frac{1}{c}\partial_t \boldsymbol{A} - \nabla\Phi + \frac{\boldsymbol{v}_s}{c} \times (\nabla \times \boldsymbol{A})\right). \tag{95}$$

With (94) and (56) we obtain

$$\boldsymbol{F}_{\text{em}} = m^*\left(\partial_t \boldsymbol{v}_s + \tfrac{1}{2}\nabla(\boldsymbol{v}_s)^2 - \boldsymbol{v}_s \times (\nabla \times \boldsymbol{v}_s)\right) = m^*(\partial_t \boldsymbol{v}_s + (\boldsymbol{v}_s\nabla)\boldsymbol{v}_s) = m^*\frac{\mathrm{d}\boldsymbol{v}_s}{\mathrm{d}t}.$$

Thus the gradient $\nabla\Phi$ of the electrostatic potential (94) does not balance the Lorentz force $\frac{1}{c}\boldsymbol{v}_s \times \boldsymbol{B}$, but both terms together account for the difference between local and substantial acceleration (70) in the superfluid. This difference is significant when, for example, the flow is stationary but does not follow a straight line.

Let us summarize how the Higgs mechanism manifests itself in the equations of motion. For a neutral superfluid with $e^* = 0$, there are many solutions to

$$\boldsymbol{J}(x) = \rho\boldsymbol{v}_s = \frac{\rho}{m^*}\left(\hbar\nabla\phi + \frac{e^*}{c}\boldsymbol{A}\right)$$

for fixed boundary conditions, corresponding to all possible configurations for the phase field $\phi(x)$. These solutions reflect the existence of the massless sound mode described by the field $\phi(x)$. For the charged superfluid, however, all the different configurations for $\phi(x)$ merely correspond to different choices of gauge; as far as the current or magnetic field distributions are concerned, all these solutions are always equivalent to one for which we have $\phi(x) = 0$. Thus the field ϕ does no longer describe an excitation. For a simply connected superconductor, it is only meaningful in that it assures gauge invariance, both on the level of the Lagrange density and on the level of the equations of motion. The solution of these equations is physically (i.e., apart from the freedom to choose the gauge) unique for a given set of boundary conditions.

We now return to the manifestation of the Higgs mechanism on the level of the Lagrangian. For a simply connected superconductor, we have already seen that we may eliminate the phase field ϕ if we introduce a new vector field A'_μ according to (77). The mass term (78) we find for A'_μ may appear to violate gauge invariance, as mass terms generally do, and may, at first sight, to be taken as a signature of a spontaneously broken gauge invariance. For one thing, however, gauge invariance is not violated. From the definition (77) it is clear that the new field simply transforms as

$$A'_\mu(x) \to A'_\mu(x)$$

under a gauge transformation (27). The Lagrangian hence remains manifestly gauge invariant, and has to remain gauge invariant, as it is the same Lagrangian as before expressed in terms of different fields. Furthermore, if a symmetry is spontaneously broken, it is never violated on the level of the Lagrangian or the Hamiltonian, but only on the level of the ground state.

In the literature, one sometimes finds the statement that "the gauge field acquires a mass" due to the Higgs mechanism. This is not exactly to the point, as it suggests that the massive vector field A'_μ is still a gauge field, while we have just seen that it is gauge invariant. In the case of a superconductor, we even know how to interpret the individual components of A'_μ physically. According to (77), (67), and (56),

$$-\frac{e^*}{c}(A'_\mu) = \left(-\frac{\mu}{c}, m^*\boldsymbol{v}_\mathrm{s}\right). \tag{96}$$

The Higgs mechanism hence does not imply that "the electromagnetic gauge field acquires a mass," but only that we can describe the superconductor in terms of gauge invariant fields, that is, in terms of the chemical potential $\mu(x)$ and the superfluid velocity $\boldsymbol{v}_\mathrm{s}(x)$. If we do this, we also have to express the Maxwell Lagrange density (34) in terms of μ and $\boldsymbol{v}_\mathrm{s}$. With

$$-\tfrac{1}{2}F_{\mu\nu}F^{\mu\nu} = \boldsymbol{E}^2 - \boldsymbol{B}^2$$

we obtain for the total Lagrange density

$$\begin{aligned}\mathscr{L} = {} & \frac{1}{8\pi e^{*2}}\left\{(\nabla\mu + m^*\partial_t\boldsymbol{v}_\mathrm{s})^2 - c^2m^{*2}(\nabla\times\boldsymbol{v}_\mathrm{s})^2\right\} \\ & - n_\mathrm{s}\left\{\mu + \left(-\mu + \frac{1}{2}m^*\boldsymbol{v}_\mathrm{s}^2\right) + \frac{1}{2m^*}\frac{1}{v^2}\left(-\mu + \frac{1}{2}m^*\boldsymbol{v}_\mathrm{s}^2\right)^2\right\}.\end{aligned} \tag{97}$$

The Euler–Lagrange equations we obtain from (97) for μ and $\boldsymbol{v}_\mathrm{s}$ are equivalent to (79)–(82), and yield exactly the same solution as above. Writing the Lagrangian in terms of μ and $\boldsymbol{v}_\mathrm{s}$ does not yield any practical advantage, but clearly illustrates that gauge invariance has become irrelevant—it is not broken, but has simply left the stage. Since all the fields are gauge invariant, (97) does not even provide a framework to think of a spontaneous violation of a gauge invariance.

These considerations apply to every field theory which displays the Higgs mechanism. In any such theory, the Lagrange density is invariant under a global physical symmetry for a matter field, and invariant under a local gauge symmetry, which affects both the matter field and the gauge field. The global symmetry is "physical" as we can classify the states of matter according to their transformation properties, while the gauge symmetry is "unphysical" as gauge transformations have no effect on the states of matter, but only on our description of these states. In our example of a superfluid, charged or neutral, the global symmetry transformation is

$$\phi(x) \to \phi(x) + \lambda, \tag{98}$$

where λ is independent of spacetime. This symmetry is spontaneously violated, which means that there are many degenerate ground states which map into each other under (98). For a neutral superfluid, we obtain a massless mode according to Goldstone's theorem. The situation is more subtle for a superconductor, as the matter field is coupled to a gauge field and the Lagrange density is also invariant under the gauge transformation (27). This "unphysical" symmetry, however, seems to contain the physical symmetry as the special case

$$\Lambda(x) = -\frac{\hbar c}{e^*}\lambda. \tag{99}$$

The formal equivalence of the transformation (98) and (27) with (99) is at the root of the widely established but incorrect interpretation of (98) as a gauge transformation, and in particular of the spontaneous violation of (98) as a spontaneous violation of a gauge symmetry. (This is presumably the reason why particle physicists like Steven Weinberg speak of "spontaneously broken gauge symmetries" interchangeably with "the Higgs mechanism.") The problem here is that the equivalence is only formal. The gauge transformation (99) represents a transformation of our description, similar to a rotation of a coordinate system we use to describe a physical state, while the transformation (98) corresponds to a transformation of our physical state, like a rotation of a physical system. Clearly, a (counterclockwise) rotation of the coordinate system has the same effect on our equations as a (clockwise) rotation of the physical system we describe with these equations, but the transformations are all but equivalent. It is not correct to refer to the spontaneous violation of (98) as a spontaneous violation of gauge symmetry. A gauge symmetry cannot be spontaneously violated as a matter of principle.

The difference between the "physical" symmetry (98) and the gauge symmetry (27) can also be appreciated at the level of conservation laws. The former yields particle number (or charge) as a conserved quantity, according to (42), while there is no conservation law associated with the latter. In the literature, (98) is often referred to as a global gauge transformation, and the conservation of charge attributed to gauge invariance. This view, however, is not consistent. If one speaks of a global gauge symmetry, this symmetry has to be a proper subgroup of the local gauge symmetry group. The alleged global gauge symmetry hence cannot be a "physical" symmetry while the local gauge symmetry is an invariance of description, or be spontaneously violated while the local symmetry is fully intact. The difference between the global phase rotation (98) and a global gauge rotation (99) is even more at evident at the level of quantum states. The BCS ground state (1) is, for example, not invariant under (98), while it is fully gauge invariant, as we have seen in Section 2.

The conclusions regarding the physical significance (or maybe better insignificance) of gauge transformations we reached here for superconductors hold for any field theory which displays the Higgs mechanism.

6. Quantum effects

This discussion of the Higgs mechanism applies only to simply connected superconductors. If we have a nontrivial topology or more than one superfluid, the phase field ϕ reassumes physical significance through its compactness, that is, the fact that its value is only defined modulo 2π. In these situations, we are not allowed to set $\phi(x) = 0$ in the equations of motion or eliminate it from the effective Lagrange density via (77) or (96), as we would loose the information regarding the compactness. Since the phase field ϕ is multiplied by Planck's constant $\hbar$ whenever it enters in the Lagrange density, any effect due to the compactness of ϕ will depend on $\hbar$, and only exist for $\hbar \neq 0$. Therefore we refer to them as "quantum effects."

The simplest of these effects in superconductors is the quantization of magnetic flux, which is analogous to the quantization of vorticity in neutral superfluids. The effect was predicted by London in a footnote in his first book [22,26] almost a decade before BCS proposed their microscopic theory. Consider a macroscopic superconductor with a hole in it, which may either be a hole in the superconducting material or a line defect or vortex in the superconducting order parameter. Like in the case of a vortex in a neutral superfluid, the phase field ϕ has to be single valued everywhere in the superconductor, but may change by a multiple of 2π as we circumvent the hole or defect along a closed curve ∂S:

$$\oint_{\partial S} \nabla\phi(x)\,\mathrm{d}\boldsymbol{l} = 2\pi n, \tag{100}$$

where n is an integer. We now take ∂S well inside the superconductor, that is, separated at each point by a distance much larger than the penetration depth λ_{L} from the hole or defect. Then, according to the Meissner effect or our derivation of London's equation above, which still applies locally, the superfluid velocity

$$\boldsymbol{v}_{\mathrm{s}} = \frac{1}{m^*}\left(\hbar\nabla\phi + \frac{e^*}{c}\boldsymbol{A}\right)$$

has to vanish along ∂S, and (100) implies

$$\oint_{\partial S} \boldsymbol{A}(x)\,\mathrm{d}\boldsymbol{l} = \int_S \boldsymbol{B}(x)\cdot\boldsymbol{n}\,\mathrm{d}a = \frac{hc}{e^*}\cdot n, \tag{101}$$

where we have used Stokes theorem once more. The magnetic flux through the hole or vortex is hence quantized in units of hc/e^*, which for $e^* = 2e$ is half of the Dirac flux quantum. Note that the vorticity (75) is not quantized in a superconductor.

We now review two further quantum effects, which are similar in neutral and charged superfluids; as in our derivation of the effective theory above, the equations for the latter case contain the former as the special case $e^* = 0$. One of the effects is phase slippage [7]. Consider two points 1 and 2 in a superfluid, which are connected by a vertical path (see Fig. 2A). Now imagine we adiabatically move a vortex from the very far left across the path to the very far right. This process yields a difference in the electrochemical potential between the two points, which is according to (68) given by

$$\Delta\mu_{\text{el.chem.}} = -\partial_t \int_1^2 \left(\hbar\nabla\phi(x) + \frac{e^*}{c}\boldsymbol{A}(x)\right)\mathrm{d}\boldsymbol{l}. \tag{102}$$

where the line integral is taken along the path between the points. The time integrated difference in the electrochemical potential is hence given by the difference

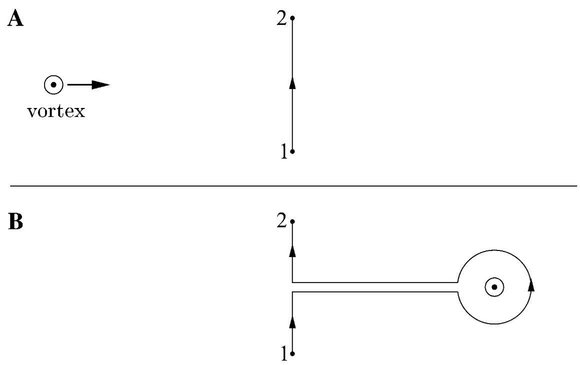


Fig. 2. Phase slippage: a vortex moving in a superfluid induces a transverse gradient in the (electro) chemical potential by dragging a branch cut in the phase of the order parameter with it.

between the line integral at the end of the process and the line integral at the beginning. Let us first consider the case of a neutral superfluid, i.e., $e^* = 0$. The line integral of $\nabla\phi$ will have changed by 2π, as the difference in the paths is topologically equivalent to encircling the vortex once (see Fig. 2B). Alternatively, we may say the vortex has dragged a branch cut across the path. (We assume that at the beginning and the end, the vortex is so far away from points 1 and 2 that we can neglect its influence on the line integral.) If we now have a continuous flow of vortices across the path, the line integral will pick up a contribution of $2\pi\hbar$ from each of them, and we obtain a chemical potential difference

$$\Delta\mu = h\langle\partial_t N_{\mathrm{v}}\rangle_{\mathrm{av}}, \tag{103}$$

where $\langle\partial_t N_{\mathrm{v}}\rangle_{\mathrm{av}}$ is the average rate of vortices crossing the path.

Let us now turn to the case of a superconductor, where we assume that during the entire process the distance between the vortex and either of the points 1 and 2 is much larger than the penetration depth. Since the line integral we obtain when encircling a superconducting vortex along a circle well inside the superconductor is zero,

$$\oint_{R\gg\lambda_{\mathrm{L}}} \left(\hbar\nabla\phi(x) + \frac{e^*}{c}\boldsymbol{A}(x)\right)\mathrm{d}\boldsymbol{l} = 0, \tag{104}$$

we do not obtain a difference in the electrochemical potential as we move an isolated vortex carrying a magnetic flux quantum across the path. So, at first sight, it may appear as there is no phase slippage effect in superconductors. The situation just described, however, is not the general one, as we dragged a unit of magnetic flux with the vortex from the very far left to the very far right. This produced a Hall effect which exactly canceled the phase slippage effect. If we consider a situation where we have a large, almost uniform magnetic field and an Abrikosov vortex lattice or liquid in which the distance between the vortices is much smaller than the penetration depth, and we have a flow of vortices across the path, the magnetic field will remain to a reasonable approximation unaffected by the flow and we recover (103) for the electrochemical potential difference. The voltage we measure between the two points is then given by

$$U = \frac{h}{2e}\langle \partial_t N_{\rm v}\rangle_{\rm av}. \tag{105}$$

This voltage is known as the Nernst effect in superconductors.

The last and possibly most striking quantum effect we review is the Josephson effect [27]. Consider two superfluids or superconductors S_1 and S_2, which are weakly coupled, say by a narrow constriction for superfluid particles or a tunneling barrier for Cooper pairs. The only requirement for the effect is that there is an energy associated with the weak link, which depends on the (gauge invariant) phase difference $\Delta\phi$ between two points 2 and 1 in superfluids S_2 and S_1:

$$E_{\rm junction} = f(\Delta\phi) \tag{106}$$

with

$$\Delta\phi \equiv \phi(2) - \phi(1) + \frac{e^*}{\hbar c}\int_1^2 \boldsymbol{A}(x)\,{\rm d}\boldsymbol{l}, \tag{107}$$

where the line integral is taken along the path the superfluid particles take when they move from one superfluid to the other. Note that $\Delta\phi$ divided by the distance between the points 2 and 1 is just the discrete version of the gradient term

$$\hbar\nabla\phi(x) + \frac{e^*}{c}\boldsymbol{A}(x)$$

we already encountered in the Ginzburg–Landau free energy, where the magnetic energy (25) was essentially given by its square. We assume that $E_{\rm junction}$ is likewise minimal for $\Delta\phi = 0$, which implies that the first term in a Taylor expansion around this minimum is quadratic in $\Delta\phi$. In the case of the junction, however, this term is not sufficient. Since ϕ is only defined modulo 2π, $f(\Delta\phi)$ has to be a periodic function of $\Delta\phi$. Josephson has shown that to a reasonable approximation, it is given by

$$f(\Delta\phi) = -E_0\cos(\Delta\phi). \tag{108}$$

Let us now assume a situation where both macroscopic superfluids are in a state of equilibrium, but the phases are not necessarily aligned relative to each other. Then only the energy stored in the junction depends on the phases of the superfluids, and the "characteristic equation" (72) becomes for superfluid S_2

$$\hbar\partial_t N_2 = \frac{\partial E_{\rm junction}(\Delta\phi)}{\partial\phi(2)}, \tag{109}$$

where N_2 is the number of superfluid particles or Cooper pairs in S_2. (We would also obtain a similar equation for S_1, but since we assume $N_1 + N_2 = \text{const.}$ and $E_{\rm junction}$ only depends on $\phi(2) - \phi(1)$, it does not provide any additional information.) The particle current from superfluid S_1 to S_2 is hence given by

$$J_{1\to 2} = \frac{1}{\hbar}\frac{\partial E_{\rm junction}(\Delta\phi)}{\partial\phi(2)} = \frac{1}{\hbar}E_0\sin(\Delta\phi). \tag{110}$$

On the other hand, since the other "characteristic equation" (66) holds for each superfluid,

$$\hbar\partial_t(\phi(2) - \phi(1)) = -(\mu(2) - \mu(1)) + e^*(\Phi(2) - \Phi(1)). \tag{111}$$

If we add

$$\partial_t \frac{e^*}{c}\int_1^2 \boldsymbol{A}(x)\,\mathrm{d}\boldsymbol{l}$$

to both sides of (111), we obtain

$$\hbar\partial_t\Delta\phi = -(\mu(2) - \mu(1)) - e^*\int_1^2 \boldsymbol{E}(x)\,\mathrm{d}\boldsymbol{l} = -\Delta\mu_{\text{el.chem.}}, \tag{112}$$

or, if we take $\Delta\mu_{\text{el.chem.}}$ time independent,

$$\Delta\phi(t) = -\frac{\Delta\mu_{\text{el.chem.}}}{\hbar}\cdot t + \Delta\phi_0.$$

Substitution into (110) yields

$$J_{1\to 2} = \frac{E_0}{\hbar}\sin(2\pi\nu t + \Delta\phi_0), \tag{113}$$

where

$$\nu \equiv -\frac{\Delta\mu_{\text{el.chem.}}}{h} \tag{114}$$

is the Josephson frequency. This implies that if the electrochemical potential is equal for both superfluids, we find a DC particle current depending on the initial alignment of the phases. If there is a difference in the potential, however, the current will oscillate with frequency ν. This is called the AC Josephson effect. The effect exists for both neutral and charged superfluids, but it is much easier to measure in a superconductor, as we can realize a difference in the electrochemical potential by applying a voltage U across the junction, $\Delta\mu_{\text{el.chem.}} = -2eU$, and easily measure oscillations in the electrical current.

Note that the Josephson effect, so astonishing its phenomenology may be, follows through the "characteristic" equations of superfluidity directly from the fact that there is a broken symmetry in superfluids and that the compact phase field which labels the different degenerate ground states is the field conjugate to the density in the superfluid. The other assumption we made in this article, the assumption that both current and momentum are carried by the same species of particles in a superfluid, was not required to explain any of the quantum effects.

Acknowledgments

I wish to thank T. Kopp, D. Schuricht, M. Vojta, and P. Wölfle for discussions of various aspects of this work.

Appendix A

In this appendix, we derive the position space wave function (4) by projecting the BCS state (1) onto a fixed number of pairs N. The (unnormalized) BCS state may be written

$$|\psi_\phi\rangle = \prod_{\boldsymbol{k}} \left(1 + \mathrm{e}^{\mathrm{i}\phi} \frac{v_{\boldsymbol{k}}}{u_{\boldsymbol{k}}} c^\dagger_{\boldsymbol{k}\uparrow} c^\dagger_{-\boldsymbol{k}\downarrow}\right)|0\rangle = \prod_{\boldsymbol{k}} \exp\left(\mathrm{e}^{\mathrm{i}\phi} \frac{v_{\boldsymbol{k}}}{u_{\boldsymbol{k}}} c^\dagger_{\boldsymbol{k}\uparrow} c^\dagger_{-\boldsymbol{k}\downarrow}\right)|0\rangle$$
$$= \exp\left(\mathrm{e}^{\mathrm{i}\phi} \sum_{\boldsymbol{k}} \frac{v_{\boldsymbol{k}}}{u_{\boldsymbol{k}}} c^\dagger_{\boldsymbol{k}\uparrow} c^\dagger_{-\boldsymbol{k}\downarrow}\right)|0\rangle = \exp\left(\mathrm{e}^{\mathrm{i}\phi} b^\dagger\right)|0\rangle.$$

The pair creation operator $b^\dagger$ is given by

$$b^\dagger \equiv \sum_{\boldsymbol{k}} \frac{v_{\boldsymbol{k}}}{u_{\boldsymbol{k}}} c^\dagger_{\boldsymbol{k}\uparrow} c^\dagger_{-\boldsymbol{k}\downarrow} = \int \mathrm{d}^3\boldsymbol{x}_1 \mathrm{d}^3\boldsymbol{x}_2 \varphi(\boldsymbol{x}_1 - \boldsymbol{x}_2) \psi^\dagger_\uparrow(\boldsymbol{x}_1) \psi^\dagger_\downarrow(\boldsymbol{x}_2)|0\rangle,$$

where $\varphi(\boldsymbol{x})$ is given by (6). If we now project out a state with N pairs according to (3), we obtain

$$|\psi_N\rangle = \int_0^{2\pi} \mathrm{d}\phi \mathrm{e}^{-\mathrm{i}N\phi} \exp\left(\mathrm{e}^{\mathrm{i}\phi} b^\dagger\right)|0\rangle = \frac{2\pi}{N!} \left(b^\dagger\right)^N |0\rangle,$$

which is (up to a normalization) equivalent to (4).

References

[1] For an excellent review and references, see S. Coleman's 1973 Erice Lectures Secret Symmetry, reprinted in S. Coleman, Aspects of Symmetry, Cambridge University Press, Cambridge, 1985.
[2] S. Weinberg, The Quantum Theory of Fields, vol. II, Cambridge University Press, Cambridge, 1996.
[3] J. Bardeen, L.N. Cooper, J.R. Schrieffer, Phys. Rev. 108 (1957) 1175.
[4] J.R. Schrieffer, Theory of Superconductivity, Benjamin, 1964.
[5] P.G. de Gennes, Superconductivity of Metals and Alloys, Benjamin, 1966.
[6] M. Tinkham, Introduction to Superconductivity, McGraw Hill, New York, 1996.
[7] P.W. Anderson, Rev. Mod. Phys. 38 (1966) 298.
[8] Note also that gauge symmetry is a local symmetry, and a local cannot break down in principle as a consequence of Elitzur's theorem;
S. Elitzur, Phys. Rev. D 12 (1975) 3978.
[9] For simplicity, we consider only time independent gauge fields and transformations.
[10] A. Auerbach, Interacting Electrons and Quantum Magnetism, Springer, Berlin, 1994.
[11] W. Marshall, Proc. R. Soc. Lond. A 232 (1955) 48.
[12] O. Penrose, L. Onsager, Phys. Rev. 104 (1956) 576;
C.N. Yang, Rev. Mod. Phys. 34 (1962) 694.
[13] P.W. Anderson, Phys. Rev. 130 (1963) 439.
[14] V.L. Ginzburg, L.D. Landau, J. Exp. Theor. Phys. 20 (1950) 1064.
[15] There are relativistic corrections to this, arising, e.g., from the binding energy of the Cooper pairs. The theory of those is still controversial;
see for example M. Liu, Phys. Rev. Lett. 81 (1998) 3223 (and references therein).
[16] R.P. Feynman, R.B. Leighton, M. Sands, The Feynman Lectures of Physics, Addison Wesley, Reading, MA, 1965;

Feynman takes the point of view that since $\Psi(x)$ may be interpreted as a condensate wave function, it should obey a Schrödinger equation. The corresponding Lagrange density is given by

$$\mathscr{L}_{\mathrm{s}}(x) = \Psi^*(x)\left\{(\mathrm{i}\hbar\partial_t + e^*\Phi(x)) - \frac{1}{2m^*}\left(-\mathrm{i}\hbar\nabla + \frac{e^*}{c}\boldsymbol{A}(x)\right)^2\right\}\Psi(x).$$

If one takes the amplitude $|\Psi(x)|$ constant, this reduces to the Lagrange density given in the text.

[17] See for example M.E. Peskin, D.V. Schroeder, An Introduction to Quantum Field Theory, Addison Wesley, Reading, MA, 1995, p. 15.

[18] Note, however, that there is no connection as of yet between the effective mass m^* in (55)–(58) and m^* in the effective Lagrange density. The equivalence is only established with (59) and (60) below.

[19] Anderson [7] refers to our definition (56) of $\boldsymbol{v}_{\mathrm{s}}(x)$ as a "pseudo-identity," as the framework of his analysis does not provide the means to recognize the approximative nature of (41).

[20] M. Greiter, F. Wilczek, E. Witten, Mod. Phys. Lett. B 3 (1989) 903.

[21] A.M.J. Schakel, Mod. Phys. Lett. B 14 (1990) 927.

[22] F. London, Superfluids, vol. I, Wiley, New York, 1950.

[23] Strictly speaking, this is London's second and more famous equation. The first equation, $\partial_t\boldsymbol{J} = -\frac{e^* n_{\mathrm{s}}}{m^*}\boldsymbol{E}$, is obtained from (68) with (65) if we neglect the $\nabla\mu$ term, which is a rather crude approximation.

[24] J. Bok, J. Klein, Phys. Rev. Lett. 20 (1968) 660.

[25] There is yet another effect which depends on the Cooper pair mass directly. If one rotates a superconductor, the rotation of the uniformly charged background induces a uniform magnetic field of strength $\boldsymbol{B} = \frac{2m^* c}{e^*}\boldsymbol{\Omega}$ well inside the superconductor, where $\boldsymbol{\Omega}$ is the angular velocity of rotation. To derive this effect, one simply has to add the electric current $e^* n_{\mathrm{s}}(\boldsymbol{\Omega}\times\boldsymbol{r})$ induced by the rotating background charge to the right of (80): $\nabla\times\boldsymbol{B} - \frac{1}{c}\partial_t\boldsymbol{E} = -\frac{4\pi e^*}{c}(\boldsymbol{J} - n_{\mathrm{s}}(\boldsymbol{\Omega}\times\boldsymbol{r}))$. Discarding the time derivative term and taking the curl, one obtains with (89) and $\nabla\times(\boldsymbol{\Omega}\times\boldsymbol{r}) = 2\boldsymbol{\Omega}$

$$\nabla\times\nabla\times\boldsymbol{B} = -\frac{4\pi e^*}{c}\left(\frac{e^* n_{\mathrm{s}}}{m^* c}\boldsymbol{B} - n_{\mathrm{s}}2\boldsymbol{\Omega}\right).$$

Solving this equation yields $\nabla\times\nabla\times\boldsymbol{B} = 0$ and therefore a uniform magnetic field deep inside the superconductor. See [22], Section 12;

J. Tate, B. Cabrera, S.B. Felch, J.T. Anderson, Phys. Rev. Lett. 162 (1989) 845.

[26] See p. 152. Note that this prediction shows that London had fully understood the significance of vector potentials in quantum mechanics almost a decade before the acclaimed work by Y. Aharonov and D. Bohm.

[27] B.D. Josephson, Phys. Lett. 1 (1962) 251.