ЖЭТФ, 2014, том 146, вып. 6 (12), стр. 1183–1191

ANDREEV BOUND STATES. SOME QUASICLASSICAL REFLECTIONS

Y. Lin*[*], *A. J. Leggett

Dept. of Physics, University of Illinois at Urbana-Champaign
61801, Urbana, USA

Received June 18, 2014

We discuss a very simple and essentially exactly solvable model problem which illustrates some nice features of Andreev bound states, namely, the trapping of a single Bogoliubov quasiparticle in a neutral s-wave BCS superfluid by a wide and shallow Zeeman trap. In the quasiclassical limit, the ground state is a doublet with a splitting which is proportional to the exponentially small amplitude for "normal" reflection by the edges of the trap. We comment briefly on a prima facie paradox concerning the continuity equation and conjecture a resolution to it.

Contribution for the JETP special issue in honor of A. F. Andreev's 75th birthday

DOI: 10.7868/S0044451014120049

1. INTRODUCTION

This year, 2014, marks not only the 75th birthday of Sasha Andreev but also the 50th anniversary of what is probably his most famous single piece of work [1], that on the reflection of an electron at a normal–superconducting boundary by conversion into a hole. Over the last half-century, the phenomenon of "Andreev reflection" has of course emerged as one of the key notions in mesoscopic physics, with applications which range far beyond the original context of the thermal conductivity of type-I superconductors in the mixed state. In this paper, we briefly discuss a "toy" problem which we feel illustrates some features of the idea in a particularly simple and intuitive way. The problem is indeed so simple that we suspect that, even if it has not been explicitly solved in the published literature in connection with a specific experimental setup, it must have been set more than once as a student exercise; nevertheless, in the present context of celebration of Sasha's work, we find it is worth a brief commentary. As a matter of history, our interest in this problem was motivated by a desire to understand whether results obtained by the standard mean-field method for some rather subtle questions concerning Berry's phase can be replicated by a strictly particle-number-conserving formalism, an issue which to our knowledge has received little discussion in the existing literature [2]. However, we do not attempt to address this issue here, and hence the level of this paper is essentially pedagogical.

Before we start, one general remark: in the original paper [1] and in much of the subsequent work on it, the phenomenon of Andreev reflection occurs as a consequence of a variation in space of the superconducting order parameter ("gap"). However, it is actually a much more general phenomenon, which crudely speaking occurs in a dense Fermi system whenever quasiparticles of a given energy are allowed in one region of coordinate space and forbidden in another, and the system is dense on the separatrix surface. This is easiest to see in the quasiclassical limit, by which we mean that all physical quantities (potential, density, order parameter, ...) are slowly varying on the scale of the mean particle separation. We consider a quasiparticle with an (initial) momentum $\mathbf{k}$ propagating from the "allowed" region towards the "forbidden" region. Since it cannot enter the latter, it must reverse its velocity. The most obvious way to do so is to reverse the momentum $\mathbf{k}$ ("normal" reflection). But it cannot do this gradually (in many small steps) because this would involve going through states of the Fermi sea which are already occupied; and it cannot do it (with any appreciable probability) in one shot, because this requires using a $q \sim 2k_F$ Fourier component of the potential (etc.) and by our definition of the quasiclassical limit any such components are ex-

[*]E-mail: yiriolin@illinois.edu

ponentially small. Hence the only option is Andreev reflection (or some analog of it in nonsuperconducting systems, cf. Ref. [3]). Of course, if we introduce abrupt spatial variations in the physical parameters, then normal reflection is no longer necessarily excluded and the situation is more complicated (cf. Ref. [4]). We make the above considerations more quantitative in Sec. 3 and the Appendix.

2. DEFINITION OF THE PROBLEM

We consider a system of $2N$ neutral fermions, initially in zero magnetic field and at $T = 0$, constrained to move in an annular container of circumference L and transverse dimensions d; for notational convenience only, we replace this geometry by a rectangular tube of length L in the z-direction and impose periodic boundary conditions in all three dimensions. (As we see in what follows, the imposition of such boundary conditions in the transverse directions is mainly a matter of convenience, but that in the longitudinal direction is crucial to our argument.) We take the fermions to interact via a short-range, spin-independent, weakly attractive potential.

We assume that the ground state $\Psi_{2N,0}$ is well approximated by the particle-conserving version of the standard BCS state, i. e., apart from normalization,

$$\Psi_{2N,0} = \left(\hat{C}^\dagger\right)^N |\text{vac}\rangle, \quad \hat{C}^\dagger \equiv \sum_k c_k a^\dagger_{k\uparrow} a^\dagger_{-k\downarrow}, \tag{2.1}$$

where $|\text{vac}\rangle$ denotes the vacuum state and the coefficients c_k are given by

$$c_k \equiv \frac{\nu_k}{u_k}, \quad \begin{matrix} u_k \\ \nu_k \end{matrix} = \frac{1}{\sqrt{2}}\left(1 \pm \frac{\varepsilon_k}{E_k}\right), \tag{2.2}$$

where $\varepsilon_k \equiv \hbar^2(k^2 - k_F^2)$ and $E_k \equiv (\varepsilon_k^2 + |\Delta|^2)^{1/2}$, with k_F being the Fermi wave vector and Δ the (isotropic) BCS energy gap, which is given by the usual self-consistent gap equation and is assumed to be $\ll E_F \equiv \hbar^2 k_F^2/2m$. The only low-energy ($E < 2\Delta$) excitations of this system are the long-wavelength density fluctuations (Anderson–Bogoliubov modes), which in the present context are of no interest to us. If we now consider the ground state and low excited states of the $(2N+1)$-particle system, these correspond to "single fermion" (Bogoliubov quasiparticle) excitations with the wave vector $\mathbf{k}$ (momentum $\hbar\mathbf{k}$), spin $\pm 1/2$, and energy E_k. The operator which, acting on the $2N$-particle ground state, creates such a Bogoliubov quasiparticle while leaving the system in a $(2N+1)$-particle number eigenstate is given by

$$\alpha^\dagger_{k\sigma} = u_k a^\dagger_{k\sigma} + \sigma \nu_k a_{-k,-\sigma} \tilde{C}^\dagger, \tag{2.3}$$

where $\tilde{C}^\dagger$ is the operator which, acting on the $2N$-particle ground state, creates the $(2N+2)$-particle ground state, i. e., apart from normalization, it is just the $\hat{C}^\dagger$ in Eq. (2.1). Although in other contexts it may be essential to remember the presence of the operator $\tilde{C}^\dagger$, it does not play a significant role in the arguments in this paper, and we mostly do not write it explicitly in what follows, simply assuming implicitly that it is always added when necessary to preserve particle number conservation. We call a Bogoliubov excitation with $\varepsilon_k > 0$ a "quasiparticle" and one with $\varepsilon_k < 0$ a "quasihole".

We now add a weak magnetic field $B(z)$ that is a function only of z and which is coupled to the spin via the Zeeman effect (only: we recall that the system is neutral!). It is convenient to take $B(z)$ to be smooth, uniformly positive, and symmetric around $z = 0$ (the middle of the tube) and to have some characteristic extension in space $R \ll L$ and characteristic magnitude B_0 which we specify below. Thus, if $\hat{H}_0$ is the original Hamiltonian of the system including (spin-independent) interactions, the complete Hamiltonian is now

$$\hat{H} = \hat{H}_0 + \sum_i \sigma_i V(z_i), \quad V(z) \equiv -\mu B(z), \tag{2.4}$$

where μ is the magnetic moment of the particles and σ_i is the projection of the spin of the ith particle on the axis of B. We now ask: What are the wave functions and energies of the ground state and low-lying energy eigenstates of the $(2N+1)$-particle system?

We can immediately say a few things. First, the ground state must certainly have a positive value of the total spin $S \equiv \sum_i \sigma_i$. Second, it must be possible to choose it to be invariant under reflection in the plane $z = 0$ (or equally under time reversal of the orbital coordinates alone: we note that we have assumed that $\hat{H}_0$ does not contain any spin–orbit interactions). Third, the qualitative behavior is intuitively obvious in the two limits of both B_0 and R large and both very small: in the former case, a substantial slab of the system becomes normal, while in the latter, the $(2N+1)$-particle ground state and low excited states correspond to spin-up single-fermion excitations which extend far beyond the region of the "trap" (the region $|z| \lesssim R$). In this paper, we are not interested in either of these limits but rather in a particular case, where intuitively speaking, in the $(2N+1)$-particle ground state, the potential $V(z)$ efficiently "traps" a single Bogoliubov quasiparticle. Moreover, we are interested in the "Ginzburg–Landau limit" in which all the relevant quantities vary slowly in space not merely over a distance k_F^{-1} but also over ξ^{-1}, where $\xi \sim \hbar v_F/\Delta$ is the Cooper-pair radius. What constraints do those requirements place on B_0 and R?

It is obvious that to be in the Ginzburg–Landau limit, we need $R \gg \xi$, and the requirement that the Zeeman coupling should not destroy superconductivity then enforces the condition $\mu B_0 \lesssim \Delta$; we shall be more conservative and require $\mu B_0 \equiv V_0 \ll \Delta$. While we should expect (in view of the $1D$ nature of the potential) a weakly bound state to exist for any B_0, the condition that it be well localized within the range of the potential, i. e., that the extra kinetic energy derived from the confinement be not too close to the binding energy, requires, as we see below, that $\hbar v_F/R \lesssim V_0 \ll \Delta$, which fortunately is already guaranteed by the condition $R \gg \xi$. Thus the necessary conditions on the parameters B_0 and R are

$$\xi \ll R \ll L, \quad V_0 \ll \Delta. \tag{2.5}$$

We note that we still have freedom to adjust the ratio

$$\alpha \equiv (V_0/\Delta)(R/\xi) \tag{2.6}$$

which essentially determines the (order of magnitude of) the number of bound states in the well, in the range ~ 1–∞.

With these conditions, $B(z)$, or equivalently $V(z) \equiv -\mu B(z)$, can be just about any smooth function: when we need a specific example we use the convenient form

$$V(z) = -V_0 \operatorname{sech}^2(z/R). \tag{2.7}$$

The Zeeman term in (2.4), with a potential $V(z)$ satisfying (2.5), is possibly the "minimal" nontrivial perturbation to the original uniform BCS problem, and as we see below allows us to derive essentially exact results for the quasiparticle spectrum within the standard mean-field approach. By contrast, the problem defined by omission of the σ_i in (2.4), i. e., that of a weak spin-independent potential, is more complicated in that in general it leads to nonzero deformation of the condensate; we do not treat it here.

3. GROUND STATE AND LOW EXCITED STATES OF THE $(2N+1)$-PARTICLE SYSTEM

Before embarking explicitly on this topic, we briefly discuss the ground state of the $2N$-particle system. In the absence of the Zeeman perturbation, this is (by hypothesis) just the simple BCS state (2.1), and it is easy to see that irrespective of the value of α in Eq. (2.6), the effect of the perturbation is at most of the order $(\mu B_0/\Delta)^2 \ll 1$. Actually, since for a completely uniform field B the effect is rigorously zero up to a critical value $B_c = \Delta/2^{1/2}\mu$, there is a strong argument that it is also rigorously zero in our problem; in any case, it is negligible within our approximations. Thus we take the $2N$-particle ground state described by Hamiltonian (2.4) as the simple uniform BCS state.

We now turn to the $(2N+1)$-particle system and discuss it from three different points of view. Unless explicitly otherwise stated, we always assume that the states we are discussing have $S = +1/2$. (States with $S = -1/2$ would certainly be unbound, and those with $|S| > 1/2$ require a minimum excitation energy close to 2Δ.) Moreover, we always assume that any states we discuss are uniform in the transverse (x, y) directions, and thus do not write these variables explicitly.

3.1. Quasiclassical approach

For the purposes of this subsection, we assume that the quantity α defined in Eq. (2.6) is large compared to unity, such that it is possible to form a quasiparticle wave packet with spread in wave vector k and position z respectively given by Δk and Δz, so as to simultaneously satisfy $\Delta k \cdot \Delta z \gg 1$ and confine the packet within the well. Then following the procedure in Ref. [5], we define the local quasiparticle energy $\tilde{E}(k, z)$ by the simple prescription

$$\begin{gathered} \tilde{E}(k,z) = E_0(k) + V(z), \\ E_0(k) \equiv (\varepsilon_k^2 + |\Delta|^2)^{1/2}. \end{gathered} \tag{3.1}$$

The quasiclassical equations of motion are then

$$\frac{dz}{dt} = \frac{\partial \tilde{E}}{\partial k}\,[k(t), z(t)] = v_F \frac{\varepsilon[k(t)]}{E_0[k(t)]} \equiv v(t), \tag{3.2a}$$

$$\frac{dk}{dt} = -\frac{1}{\hbar}\frac{dV}{dz}[z(t)], \tag{3.2b}$$

where $v_F = d\varepsilon/dk|_{k=k_F} = \hbar k_F/m$. We see that $v(t)$ is in the positive (negative) direction for $k(t) > k_F$ ($k(t) < k_F$). From Eq. (3.2), it follows immediately that the quantity $\tilde{E}(k, z)$ defined in Eq. (3.1) is a constant of motion:

$$\tilde{E}\,(k(t), z(t)) = \text{const} \equiv \tilde{E} \tag{3.3}$$

and hence the wave vector $k(t)$ is given, in the limit $|\varepsilon_k| \ll \epsilon_F$ which is of most interest to us, by the expression

$$k(t) = \\ = \pm\left(k_F \pm \frac{1}{\hbar v_F}\left(\left(\tilde{E}-V[z(t)]\right)^2 - |\Delta|^2\right)^{1/2}\right). \quad (3.4)$$

The motion described by Eq. (3.2) could hardly be simpler. We suppose that we start with a wave packet which has (approximately) $z(0) = 0$ and $k(0) > k_F$. For $t > 0$, this packet moves rightwards, gradually decreasing $k(t)$ according to Eq. (3.2b) and correspondingly $v(t)$ according to (3.2a), until it reaches the point $z_c(E)$ defined by

$$W(z_c) = \tilde{E} - \Delta, \quad (3.5)$$

at which $k(t) = k_F$. At this point, according to Eq. (3.2a), its velocity vanishes and as $k(t)$ passes through k_F, it converts itself into a leftward-traveling hole, $\varepsilon_k < 0$. At the left-hand turning point, at which Eq. (3.5) is again satisfied, the inverse process takes place: the left-moving quasihole is converted back into a right-moving quasiparticle. In the limit $V_0 \ll \Delta$ ("shallow" well), the period of the cyclic motion is

$$T = 2\int_{-z_c}^{z_c} \frac{dz}{v(z)} = \\ = \frac{2}{v_F}\int_{-z_c}^{z_c} \frac{\left(\tilde{E} - V(z)\right) dz}{\left(\left(\tilde{E} - V(z)\right)^2 - |\Delta|^2\right)^{1/2}} \approx \\ \approx \frac{(2\Delta)^{1/2}}{v_F}\int_{-z_c}^{z_c} \frac{dz}{\left(-\left(\delta + V(z)\right)\right)^{1/2}}, \quad (3.6)$$

where $\delta \equiv \Delta - \tilde{E} (> 0)$, and the last approximate equality holds for $\delta \ll \Delta$. Comparing expression (3.6) with the standard expression for the period of a single particle of mass m moving in the potential $V(z)$ with the total energy η in the absence of the Fermi sea, namely,

$$T \equiv (2m)^{1/2}\int_{-z_c}^{z_c} \frac{dz}{\left(\eta - V(z)\right)^{1/2}}, \quad (3.7)$$

it is tempting to define an "effective mass" m^* by

$$m^* \equiv \Delta/v_F^2 \left(\sim (\Delta/E_F)m\right). \quad (3.8)$$

However, it should be remembered that this is not the ratio of momentum to velocity (which actually changes sign over the course of the cycle).

An important point to note in this quasiclassical approach is that in view of the invariance of the Hamiltonian under time reversal of the orbital coordinates alone (irrespective of whether the potential $V(z)$ is symmetric), any cyclic motion of the type described above has a time-reversed partner, in which a left-moving quasiparticle converts (at the left-hand turning point) into a right-moving quasihole. (Formally, this is achieved by the substitution $k(t) \to -k(t)$.)

3.2. Ground state in the quasiclassical approximation

For the simplest $(2N + 1)$-particle energy eigenstates, including the ground state, we write the standard mean-field ansatz

$$\Psi_{2N+1} = \\ = \int dz\left(u(z)\hat{\psi}_\uparrow^\dagger(z) + \nu^*(z)\hat{\psi}_\downarrow(z)\tilde{C}^\dagger\right)\Psi_{2N,0}, \quad (3.9)$$

where $\hat{\psi}_\uparrow^\dagger(z)$ ($\hat{\psi}_\downarrow(z)$) is the standard Fermi creation (annihilation) operator for a spin-up (spin-down) particle, and the "particle" component $u(z)$ and the "hole" component $v(z)$ of the Bogoliubov quasiparticle wave function obey the standard Bogoliubov–de Gennes (BdG) equations. We need to remember that the Zeeman energy of a spin-down hole is the same as that of a spin-up particle (unlike the case of a spin-independent potential), and that the creation of an extra Cooper pair (the $\tilde{C}^\dagger$ in Eq. (3.9)) costs an energy 2μ, where within the usual BCS approximation we can identify the chemical potential μ with ϵ_F. Thus the correct form of the BdG equations for $S = +1/2$ is

$$\left\{-\frac{\hbar^2}{2m}\frac{d^2}{dz^2} + V(z)\right\}u(z) + \Delta v(z) = \\ = (\epsilon_F + \Delta + E)\,u(z), \quad (3.10a)$$

$$\Delta u(z) + \left\{\frac{\hbar^2}{2m}\frac{d^2}{dz^2} + 2\epsilon_F + V(z)\right\}u(z) = \\ = (\epsilon_F + \Delta + E)\,v(z), \quad (3.10b)$$

where for subsequent convenience (in this subsection only) we have taken the zero of energy E at $\epsilon_F + \Delta$, the minimum excitation energy needed to add an extra fermion in the absence of the Zeeman potential. Making the standard substitution [1]

$$u(z) = \exp(ik_F z)f(z), \\ v(z) = \exp(ik_F z)g(z), \quad (3.11)$$

discarding the terms in $d^2/dz^2(f(z), g(z))$ (we return to this point below), and combining equations (3.10a) and (3.10b), we obtain the single equation

$$(\hbar^2 v_F^2)\frac{d^2 f}{dz^2} + \left\{(E+\Delta - V(z))^2 - \Delta^2\right\} f(z) + + i\hbar v_F \frac{dV(z)}{dz} f(z) = 0. \quad (3.12)$$

In the case of interest, $V_0 \ll \Delta$, not only the condition $|V(z)| \ll \Delta$ but also, for any bound state, $|E| \ll \Delta$, must be fulfilled, as a result of which Eq. (3.12) simplifies to the linear eigenvalue equation

$$\frac{\hbar^2}{2m^*}\frac{d^2}{dz^2} f + (E - V(z))\, f(z) + i\xi \frac{dV}{dz} f(z) = 0, \quad (3.13)$$

where, as in Sec. 1, we define $m^* \equiv \Delta/v_F^2$ and also a quantity $\xi \equiv \hbar v_F/\Delta$, which up to a numerical factor is the Cooper pair radius (or the Pippard coherence length). The equation for $g(z)$ is the same except that the sign of the last term is reversed. We note that if we had chosen to take out a factor $\exp(-ik_F z)$ rather than $\exp(+ik_F z)$ in Eq. (3.11), the only effect would have been to change the sign of the last term in Eq. (3.13) and the corresponding equation for $g(z)$.

If we temporarily neglect the last term in Eq. (3.13), the resulting equation is exactly the standard time-independent Schrödinger equation (TISE), and moreover, in the case of a bound state, must satisfy the standard boundary condition $f(z) \to 0$ for $|z| \to \infty$. Consequently, we can apply all the standard textbook lore for the solution of the TISE, and the original components $u(z)$ and $v(z)$ of the BdG spinor are given (up to normalization) by

$$u(z) = \exp(ik_F z)\psi_{Sch}(z), \quad (3.14a)$$

$$v(z) = \exp(ik_F z)\left\{i\xi\frac{d}{dz} + \frac{\Delta + E - V(z)}{\Delta}\right\} \times \times \psi_{Sch}(z) \approx u(z), \quad (3.14b)$$

where $\psi_{Sch}(z)$ is the solution of the Schrödinger equation for a particle of the mass $m^* \equiv \Delta/v_F^2$ in the potential $V(z)$, and the approximate equality in Eq. (3.14b) holds in the limit of interest $V_0 \ll \Delta$ and $\xi \ll R$.

It is interesting at this point to estimate the order of magnitude of the energy splittings ΔE and the spatial extent Δz of the low-lying bound states. According to Eq. (3.8), these are respectively $\sim (E_F/\Delta)^{1/2}$ and $(E_F/\Delta)^{1/4}$ times the values they would take for a free particle in the potential $V(z)$. Explicitly,

$$\Delta E \sim (\hbar^2 V_0/m^* R^2)^{1/2} \sim (\xi/R)(V_0\Delta)^{1/2}, \quad (3.15a)$$

$$\Delta z \sim (\hbar R)^{1/2}(m^* V_0)^{-1/4} \sim \sim R(\Delta\xi^2/V_0 R^2)^{1/4}. \quad (3.15b)$$

From (3.15a), we see that the last term in Eq. (3.13) is of the order $(V_0/\Delta)^{1/2}$ relative to ΔE, and hence at first glance it can be treated perturbatively and has no major qualitative effect. We also note that the terms in $d^2/dz^2(f,g)$ which we omitted in deriving Eq. (3.12) are of the order $(k_F\Delta z)^{-1} \ll 1$ relative to those kept.

However, a very simple argument suffices to show that the many-body wave function given by inserting (3.14a)–(3.14b) in (3.9) cannot be even approximately the $(2N+1)$-particle ground state (or more generally that functions corresponding to (3.14) cannot be even approximate energy eigenstates). For were it so, then by the continuity equation the divergence of the particle current $j(r)$ would necessarily be zero. However, $j(r)$ is manifestly nonzero in the region of the Zeeman trap, while far from the latter, both components of the BdG wave function vanish and all that is left is the Cooper pairs, which because of the quantization condition carry no current. Hence, states corresponding to the forms (3.14) of $u(z)$ and $v(z)$ can never be energy eigenstates.

Unsurprisingly, the immediate resolution of the difficulty lies in the recognition that since the "orbital time-reversal operator" $\hat{T}_{orb}$ (defined by $\hat{T}_{orb}\hat{\sigma}_i\hat{T}_{orb}^{-1} = \hat{\sigma}_i$, $\hat{T}_{orb}\hat{x}_i\hat{T}_{orb}^{-1} = \hat{x}_i$, and $\hat{T}_{orb}\hat{p}_i\hat{T}_{orb}^{-1} = -\hat{p}_i$) commutes with the Hamiltonian in (2.4), and states of the form (3.13) are not eigenstates of $\hat{T}_{orb}$, and each state of this form is accompanied by a time-reversed partner that is exactly degenerate with it in the approximations made below Eq. (3.15). Thus, even if the mixing of time-reversal states is exponentially small (as we see below), it is nevertheless necessary to take it fully into account.

If we return to our original BdG equations (3.10) and let $\Psi_{R,L}$ denote the 2-component spinors (u,v) corresponding to the choice $\exp(\pm ik_F z)$ in (3.11), then the associated mean-field (Nambu) Hamiltonian H_{BdG} includes not only the terms diagonal in the (R,L) basis, which we have dealt with and which lead to solutions (3.14), but also (Hermitian) off-diagonal terms $\langle\Psi_R|H_{BdG}|\Psi_L\rangle = \langle\Psi_L|H_{BdG}|\Psi_R\rangle^*$. Physically, these terms correspond to the reflection of a Bogoliubov quasiparticle by the edges of the Zeeman potential in a "normal" $k_F \to -k_F$ process rather than an Andreev one; this point comes out rather more clearly in the formulation in Sec. 3.3. However, we should expect these off-diagonal terms to be of the order of the quantity

$$K_0 \equiv \int \exp(2ik_F z)V(z)|f(z)|^2 dz. \quad (3.16)$$

Using the estimate (3.15b) for the range of $f(z)$ and the assumed smoothness of $V(z)$, we find

$$|K_0| \sim \exp(-\zeta), \quad \zeta \sim k_F R(\Delta\xi^2/V_0 R^2)^{1/4} \gg 1. \tag{3.17}$$

In view of the exact degeneracy of the original states of Ψ_R and Ψ_L, the effect of the exponentially small off-diagonal term K_0 is to make the true $(2N+1)$-particle energy eigenstates the linear combinations

$$\Psi_\pm = 2^{-1/2}(\Psi_L \pm \Psi_R) \tag{3.18}$$

with an energy splitting $2|K_0|$. States (3.18) are eigenstates of the orbital time reversal operator $\hat{T}_{orb}$, as they should be.

It should be noted, however, that this maneuver does not totally remove the paradox raised above (see Sec. 4).

3.3. An alternative approach beyond the quasiclassical approximation

In this subsection, we discuss the Andreev bound states in momentum space, where the effect of the external Zeeman potential on the formation of the quasiparticle bound states is particularly manifest, especially concerning normal reflection processes that govern the energy splitting of the ground state doublet (see details below). In this approach, we do not assume the quasiclassical approximation as in the above two subsections, and the resulting equation of motion for an Andreev state contains spatial variations in all length scales. We then discuss physical features of the bound solutions to the equation of motion. We start by expanding the $(2N+1)$-particle energy eigenstates in terms of plane-wave BdG quasiparticle states

$$\psi_{2N+1} = \sum_k C_k \alpha^\dagger_{k\uparrow}\psi_{2N,0}, \tag{3.19}$$

where $\alpha^\dagger_{k\uparrow}$ are BdG quasiparticle creation operators with the momentum k and spin $+1/2$ (we restrict ourselves to spin $+1/2$ excitations), and the C_k (not to be confused with the c_k in Eq. (2.1)) are complex coefficients. In the absence of the external Zeeman field, each plane-wave BdG quasiparticle generates a $(2N+1)$-particle energy eigenstate in a homogeneous superconducting system. The lowest-energy eigenstate has a quasiparticle momentum at the Fermi momentum. In the presence of the Zeeman field, we can form localized wave packets inside the trap to take advantage of the Zeeman energy. Because a localized wave packet is composed of plane-wave quasiparticle states with different momenta, there is an associated energy spread which we call the "kinetic energy" for convenience. Therefore, the problem looks like the standard single-particle Schrödinger-like problem, and the energy ground state is achieved by minimizing the sum of kinetic and potential (Zeeman) energies.

We now make the argument more quantitative by writing the Zeeman energy in terms of plane-wave quasiparticle operators. In the second-quantized form, the Zeeman Hamiltonian is given by

$$H_Z = -\sum_{k,k'} V_{k-k'}\left(\alpha^\dagger_{k'\uparrow}\alpha_{k\uparrow} - \alpha^\dagger_{k'\downarrow}\alpha_{k\downarrow}\right). \tag{3.20}$$

We then expand electron operators in terms of BdG quasiparticle operators and rewrite (3.20) in terms of BdG quasiparticles:

$$H_Z = -\sum_{k,k'} V_{k-k'}\Big((u_{k'}\alpha^\dagger_{k'\uparrow} - \nu^*_{k'}\alpha_{-k'\downarrow}) \times (u_k\alpha_{k\uparrow} - \nu_k\alpha^\dagger_{-k\downarrow}) - (\nu^*_{-k'}\alpha_{-k'\uparrow} + u_{-k'}\alpha^\dagger_{k'\downarrow}) \times (\nu_{-k}\alpha^\dagger_{-k\uparrow} + u_{-k}\alpha_{k\downarrow})\Big). \tag{3.21}$$

The relevant terms for bound states are

$$H_{Z,eff} = -\sum_{k,k'} V_{k-k^2}(u_{k'}u_k\alpha^\dagger_{k'\uparrow}\alpha_{k\uparrow} + \nu^*_{-k'}\nu_{-k}\alpha^\dagger_{-k\uparrow}\alpha_{-k'\uparrow}). \tag{3.22}$$

Terms in (3.21) involving two quasiparticle annihilators annihilate $(2N+1)$-particle states and those creating two quasiparticles generate states with the energy higher by the energy gap than the relevant energy range we are interested in. Therefore, the only relevant terms are those that scatter plane-wave quasiparticle states from one momentum to another, just as the usual potential well scatters a single particle.

Equation (3.22) together with the BCS Hamiltonian yields the total Hamiltonian

$$H = \sum_k E_k\alpha^\dagger_{k\uparrow}\alpha_{k\uparrow} - \sum_{k,k'} V_{k-k'}(u_{k'}u_k + \nu_{k'}\nu^*_k)\alpha^\dagger_{k'\uparrow}\alpha_{k\uparrow}, \tag{3.23}$$

where we have only included spin-up quasiparticles as spin-down quasiparticles are not bound by the Zeeman field and are irrelevant to the discussion. Now the problem reduces to finding C_k in Eq. (3.19) to diagonalize the Hamiltonian (3.23). This is very much like the Schrödinger problem for a single particle in a potential well, with two exceptions. First, the kinetic energy is $E_k = \sqrt{\varepsilon_k^2 + \Delta^2}$ instead of $\varepsilon_k = \hbar^2k^2/2m - \mu$, and hence the minimum kinetic energy is achieved at $\pm p_F$, instead of at zero. This has important consequences on the energy spectrum, which we discuss shortly. Second, the potential contains coherence factors and is not the standard external potential in that it depends on k and k' separately. Given that the Zeeman field is wide compared to the Cooper pair radius, the difference between k and k' for which the Fourier transform of the Zeeman field is appreciable is small such that $|\varepsilon_k - \varepsilon_{k'}|$ is small compared to the gap Δ, and we can therefore approximate the coherence factors by 1. We finally arrive at the equation of motion for Andreev bound states

$$E_k C_k - \sum_{k'} V_{k-k'} C_{k'} = E C_k. \tag{3.24}$$

By expanding E_k around $k = \pm k_F$ to the leading order and extracting a phase factor $\exp(\pm i k_F z)$, we recover Eq. (3.13) without the last spatial gradient term. If we take the expectation energy of the $(2N+1)$-particle state (3.19), we obtain the kinetic energy $\sum_k E_k |C_k|^2$ and the potential energy $-V_{k-k'} C_k C^*_{k'}$. It is clear that we can save potential energy by including C_k from both around k_F and $-k_F$. The energy splitting is however exponentially small since $V_{k-k'}$ decays exponentially as $\exp(-k_F \Delta z)$, which is consistent with the conclusion in the last subsection. It is interesting to note that we can apply the same analysis to the standard Andreev setup, where it is the gap that is varying in space. Basically, all we need to do is to Fourier transform the gap function to the momentum space and write the mean-field particle–particle interactions in terms of plane-wave BdG quasiparticles, and the rest of the analysis follows. We see that the gap plays the similar role as ordinary potential in trapping the Andreev bound states. In particular, the energy splitting is exponentially small for the low bound states in a slowly varying gap.

4. PROBLEM CONCERNING THE CONTINUITY EQUATION

We first return to the considerations raised in Sec. 3.2 concerning the continuity equation[1)]. We first note that in the conventional (particle-number-nonconserving) approach to the BdG equations, there is, at least prima facie, no paradox even in the limit of complete neglect of normal reflection processes, since the lack of particle number conservation in that approach permits the continuity equation to be violated. In our particle-number-conserving approach, however, that loophole is closed and, moreover, the "resolution" given in Sec. 3.2 is not truly satisfactory, since over timescales much shorter than the exponentially long lifetime against normal reflection the system, even if it is in fact, e. g., in state (3.14), should not "know" that it is not in an energy eigenstate, and thus, for the expectation value in the state (3.14) we should have $\langle \operatorname{div} j(r)\rangle = -(\partial/\partial t)\langle \rho(r)\rangle = 0$.

We believe, therefore, that the following conclusion is inescapable: in an accurate description of an Andreev bound state, any motion of the Bogoliubov quasiparticle must inevitably be accompanied by a deformation of the condensate, which is such as to guarantee that the divergence of the total current be zero. Since this latter condition, while necessary for the resultant state to be an energy eigenstate, is not sufficient, it does not by itself suffice to identify the appropriate deformation. However, it is energetically advantageous for the system to have zero current since in order for the kinetic energy cost in the condensate due to the deformation to be finite, the current has to go to zero in the thermodynamic limit. In the $3D$ thermodynamic limit (both $L, d \to \infty$), the energy necessary to form the deformation tends to zero as $1/d^2$, simply because the quasiparticle probability density itself has this dependence. However, in the "$1D$ thermodynamic limit" ($L \to \infty$ and $d \to$ const), the energy tends to a constant. But the energy saving of the quasiparticle due to the flow of the condensate (the "$\mathbf{v}\cdot\mathbf{p}$" term) is of the same order of magnitude as that of the energy cost and is likely to cancel the latter. Hence, the expectation energy of the system with a deformed condensate is as low as the state without the deformation. Thus we are inclined to believe that the appropriate deformation of the condensate when the Bogoliubov quasiparticle is in the approximate energy eigenstate (3.14) consists in multiplying the $2N$-particle ground state (2.1) by the function $\exp\left(-i\sum_i f(z_i)\right)$, where

$$f(z) = \frac{k_F L}{2N} \int |\psi_{Sch}(z)|^2 dz. \tag{4.1}$$

If this is correct (and indeed more generally), then the true states of the ground state doublet involve a small but nonzero entanglement between the states of the Bogoliubov quasiparticle and that of the condensate. However, we have not at the time of writing established that the conjecture represented by Eq. (4.1) in fact constitutes the ground state of the $(2N+1)$-particle many-body problem in some well-defined "post-BdG" approximation.

[1)] The considerations in this section should be regarded as rather preliminary and possibly subject to revision, particularly in view of the possible connection to papers [6, 7] whose relevance we appreciated only at a very late stage in the preparation of this manuscript.

This material is based on work supported by the National Science Foundation under Award No. NSF-DMR-0906921. It is a pleasure to dedicate this paper to Sasha Andreev and to wish him many more happy and fruitful years of research in physics.

APPENDIX

We discuss the case where the external Zeeman potential has the shape of a square well. In this situation, the standard quasiclassical approach for the slowly varying potential is not applicable. We discuss the problem in the standard textbook approach by matching boundary conditions for the wave functions at the trap edges. We note that in this appendix, as in Secs. 3.1 and 3.3 but differently from Sec. 3.2, the zero of quasiparticle energy is the conventional one (the chemical potential μ).

The BdG equation can be written as

$$\begin{aligned} \left(\frac{p^2}{2m} - \mu\right) u + \Delta \nu = (E+V)u, \\ -\left(\frac{p^2}{2m} - \mu\right) \nu + \Delta u = (E+V)\nu, \end{aligned} \qquad (A.1)$$

where the potential V is constant in the trap and zero outside. The trap extends from $z = -W/2$ to $z = W/2$. We use the momentum operator p here in the BdG Hamiltonian to emphasize that a solution is superposition of plane waves, each of which satisfies the BdG equation either inside or outside the trap. To match the boundary conditions at the trap edges, we need to superpose these plane-wave states to obtain true eigenstates. It is convenient to define ratio of u and ν by

$$\begin{aligned} \frac{\nu}{u} = \frac{\Delta}{E+\Omega_o^{\pm}} = F_o^{\pm}, \\ \frac{\nu}{u} = \frac{\Delta}{E+V+\Omega_i^{\pm}} = F_i^{\pm}, \end{aligned} \qquad (A.2)$$

where the subscripts o and i respectively refer to the outside and inside of the Zeeman trap, with $\Omega_o^{\pm} = \pm i\sqrt{\Delta^2 - E^2}$ and $\Omega_i^{\pm} = \pm\sqrt{(E+V)^2 - \Delta^2}$. We note that $\Omega_o^{\pm}$ is pure imaginary and this is the factor which gives rise to the exponential decay of the bound solutions outside the trap. Strictly speaking, there are eight boundary conditions, four of them coming from continuity conditions for u and v at the two trap edges, the other half coming from the continuity conditions for the first derivatives of u and v. Upon Andreev reflection, the momentum change is of the order of Δ/E_F compared to the Fermi momentum, and therefore the momenta of the particle and hole plane-wave solutions differ only by the order Δ/E_F, relative to the Fermi momentum. If we ignore this small difference, the plane-wave solutions for wave vectors in opposite directions become separate. We only need to match the continuity conditions for u and v (with wave vectors in one direction) and the continuity conditions for the first derivatives are automatically satisfied (this point is discussed in [8]). Within this approximation, the solutions with momentum in two opposite directions are degenerate, as is the case for the general gradually varying potential discussed in the main text. We consider a solution inside the trap of the form

$$\begin{aligned} u_i &= u_i^+ e^{ip_i^+ z} + u_i^- e^{ip_i^- z}, \\ \nu_i &= F_i^+ u_i^+ e^{ip_i^+ z} + F_i^- u_i^- e^{ip_i^- z}, \end{aligned} \qquad (A.3)$$

where $p_i^{\pm} = p_F + \Omega_i^{\pm}/\hbar v_F$. Similarly, the solution outside the trap is given by

$$\begin{aligned} u_o &= u_o^+ e^{ip_o^+ z}\Theta(z) + u_o^- e^{ip_o^- z}\Theta(-z), \\ \nu_o &= F_o^+ u_o^+ e^{ip_o^+ z}\Theta(z) + F_o^- u_o^- e^{ip_o^- z}\Theta(-z), \end{aligned} \qquad (A.4)$$

where $p_0^{\pm} = p_F + \Omega_o^{\pm}/\hbar v_F$, and $\Theta(z) = 1$ for $z > 0$ and $\Theta(z) = 0$ for $z < 0$.

Now by matching the boundary conditions at $z = W/2$ and $z = -W/2$, we obtain the following equations at $z = W/2$:

$$\begin{aligned} u_o^+ e^{ip_o^+ W/2} &= u_i^+ e^{ip_i^+ W/2} + u_i^- e^{ip_i^- W/2}, \\ F_o^+ u_o^+ e^{ip_o^+ W/2} &= F_i^+ u_i^+ e^{ip_i^+ W/2} + \\ &\quad + F_i^- u_i^- e^{ip_i^- W/2}. \end{aligned} \qquad (A.5)$$

Similarly, the equations at $z = -W/2$ are

$$\begin{aligned} u_o^- e^{-ip_o^- W} &= u_i^+ e^{-ip_i^+ W/2} + u_i^- e^{-ip_i^- W/2}, \\ F_o^- u_o^- e^{-ip_o^- W/2} &= F_i^+ u_i^+ e^{-ip_i^+ W/2} + \\ &\quad + F_i^- u_i^- e^{-ip_i^- W/2}. \end{aligned} \qquad (A.6)$$

Combining Eq. (A.6) with Eq. (A.5), we obtain the equation

$$\frac{F_i^- - F_o^-}{F_o^- - F_i^+} = \frac{F_i^- - F_o^+}{F_o^+ - F_i^+} e^{i(p_i^- - p_i^+)W}. \qquad (A.7)$$

It can be written as

$$\begin{aligned} &\mathrm{tg}^{-1}\left(\frac{\sqrt{\Delta^2 - E^2}}{V - \sqrt{(E+V)^2 - \Delta^2}}\right) - \\ &\quad - \mathrm{tg}^{-1}\left(\frac{\sqrt{\Delta^2 - E^2}}{V + \sqrt{(E+V)^2 - \Delta^2}}\right) = \\ &\quad = -\frac{\sqrt{(E+V)^2 - \Delta^2}}{\hbar v_F} W + n\pi, \end{aligned} \qquad (A.8)$$

where n is a positive integer.

We can find simple solutions of (A.8), in the wide trap limit, i. e., satisfying the condition

$$(E+V)^2 - \Delta^2 \ll V^2. \tag{A.9}$$

This condition is equivalent to

$$E + V - \Delta \sim \frac{\hbar^2 v_F^2}{\Delta W^2} \ll \frac{V^2}{\Delta}, \quad \Delta W \ll L_\nu, \tag{A.10}$$

where $L_\nu = \hbar v_F / V$ is the length scale associated with the trap strength.

Under this condition, we can expand Eq. (A.8) to the first order in $\sqrt{(E+V)^2 - \Delta^2}/V$, with the result

$$\frac{1}{1 + \dfrac{\Delta^2 - E^2}{V^2}} = -\frac{WV}{2\hbar v_F} + \frac{n\pi}{2\sqrt{\dfrac{(E+V)^2-\Delta^2}{V^2}}}. \tag{A.11}$$

We now consider low-energy bound states such that $(E+V-\Delta)/V \ll 1$ and $(\Delta - E)/V \sim 1$. Since the LHS of Eq. (A.11) is much smaller than 1 and the absolute values of both terms in the RHS of Eq. (A.11) are much greater than 1, we can set the LHS to zero. Hence, we arrive at the solution

$$E = \sqrt{\left(\frac{n\pi\hbar v_F}{W}\right)^2 + \Delta^2} - V \approx \Delta - V + \frac{n^2\hbar^2\pi^2}{m^* W^2}, \tag{A.12}$$

where m^* is the effective mass introduced in Eq. (3.8).

This solution is consistent with the intuitive argument. For low bound states, all Zeeman energy V is saved since low bound wave functions are completely localized inside the trap, i. e., its range outside the trap is negligible, hence the term $-V$ in Eq. (A.12). The term $(n\pi\hbar v_F/W)^2$ in the above equation is simply the kinetic energy of the quasiparticle since its momentum $\delta p = p - p_F$ is quantized by the trap as $2n\hbar\pi/(2W)$.

It is interesting to compare this result with the spectrum for a quasiparticle trapped between two superconductors in an SNS junction [9]. There, we have a square-well form of the gap. Inside the well, the quasiparticle is in the superposition of a normal particle and a normal hole, and the effective quantization length is twice the trap width W due to Andreev reflection, and hence the spectrum is

$$E = \frac{2n\pi\hbar v_F}{2W}. \tag{A.13}$$

Our result (A.12) is consistent with the standard picture that due to the Andreev reflection, the quantization length is twice the trap width.

We have so far neglected the mixture of plane-wave solutions with wave vectors in opposite directions. In order to see the energy splitting between these time-reversal states, we need to include wave vectors in both directions and match boundary conditions for both wave functions and their gradients. The calculation is much more complicated and the eigenvalue equations are rather lengthy and are not listed here; we merely note (a) that neglecting terms of the order $\sqrt{(\hbar^2/m^*W^2)/V}$ relative to the leading term in the eigenvalue equations, we recover the result in (A.12), and (b) that the term which leads to the splitting of the doublets is of the order $1/k_F W$ relative to the level separation in (A.12), indicating that the effect of normal reflection falls off only as an inverse power law in the "semiclassical" parameter $k_F W$, not exponentially as in the case of the "smooth" well discussed in the main text. This is, of course, what we would expect in the light of existing results such as those in Ref. [4].

REFERENCES

1. A. F. Andreev, Zh. Eksp. Teor. Fiz. **46**, 1823 (1964).
2. A. J. Leggett and Yiruo Lin, in: *In Memory of Akira Tonomura*, ed. by Kazuo Fujikawa and Yoshimasa A. Ono, World Scientific, Singapore (2014), p. 74.
3. B. Wang, J. Peng, D. Y. Xing, and J. Wang, Phys. Rev. Lett. **95**, 086608 (2005).
4. G. E. Blonder, M. Tinkham, and T. M. Klapwijk, Phys. Rev. B **25**, 4515 (1982).
5. N. A. Greaves and A. J. Leggett, J. Phys. C: Sol. State Phys. **16**, 4383 (1983).
6. J. Sanchez-Canizares and F. Sols, J. Phys.: Condens. Matter **8**, L207 (1996).
7. J. Sanchez-Canizares and F. Sols, Physica B **252**, 304 (1998).
8. J. Demers and A. Griffin, Can. J. Phys. **49**, 285 (1971).
9. I. O. Kulik, JETP **30**, 944 (1970).