Can One Determine the Underlying Fermi Surface in the Superconducting State of Strongly Correlated Systems?

Rajdeep Sensarma, Mohit Randeria, and Nandini Trivedi
Department of Physics, The Ohio State University, Columbus, Ohio 43210, USA
(Received 30 June 2006; published 8 January 2007)

The question of determining the underlying Fermi surface (FS) that is gapped by superconductivity (SC) is of central importance in strongly correlated systems, particularly in view of angle-resolved photoemission experiments. Here we explore various definitions of the FS in the superconducting state using the zero-energy Green’s function, the excitation spectrum, and the momentum distribution. We examine (a) \(d\)-wave SC in high-\(T_c\) cuprates, and (b) the \(s\)-wave superfluid in the BCS–Bose-Einstein condensation (BEC) crossover. In each case we show that the various definitions agree, to a large extent, but all of them violate the Luttinger sum rule and do not enclose the total electron density. We discuss the important role of chemical potential renormalization and incoherent spectral weight in this violation.

DOI: 10.1103/PhysRevLett.98.027004 PACS numbers: 74.25.Jb, 74.20.Fg, 74.72.-h, 74.90.+n

The Fermi surface (FS), the locus of gapless electronic excitations in \(k\)-space, is one of the central concepts in theory of Fermi systems. In a Landau Fermi liquid at \(T = 0\), Luttinger [1] defined the FS in terms of the single-particle Green’s function \(G^{-1}(k, 0) = 0\) and showed that it encloses the same volume as in the noninteracting system, equal to the fermion density \(n\). In many Fermi systems of interest the ground state has a broken symmetry. Here we study states with superconducting (SC) long-range order, where there is no surface of gapless excitations, and ask the following question: Is there any way to define at \(T = 0\) the “underlying Fermi surface” that got gapped out by superconductivity?

From a theoretical point of view, this question is of relevance to all superconductors irrespective of pairing symmetry or mechanism. The answers turn out to be of particular interest for strongly correlated superconductors, where the surprising effects that we find are large enough to be measured experimentally. Angle-resolved photoemission spectroscopy (ARPES) [2] has emerged as one of the most powerful probes of complex materials and has been extensively used to determine the FS in strongly correlated systems, often from data in the SC state [3,4]. One of our goals is to understand exactly what a \(T = 0\) measurement can tell us about the FS. This is especially important in the cuprates where the normal state must necessarily be studied at high temperatures and does not show sharp electronic excitations, expected in Fermi liquids, in contrast to the SC state which does show sharp Bogoliubov quasiparticles. Our results are also of interest for a completely different class of systems: strongly interacting Fermi atoms [5] in the BCS–Bose-Einstein condensation (BEC) crossover [6,7]. Here too the question of an underlying Fermi surface is of direct experimental relevance [8].

In this Letter, we first show that Luttinger’s original argument [1] cannot be generalized to the SC state, and this violation is related to broken gauge invariance [9]. We then explore various criteria for defining the “underlying Fermi surface” in the \(T = 0\) SC state, using properties of the single-particle Green’s function \(G(k, \omega)\) [10] directly related to experimentally measurable quantities. We present results for the two systems described above: (a) the \(d\)-wave SC state of the high-\(T_c\) cuprates which is dominated by strong Coulomb correlations, and (b) the \(s\)-wave superfluid state in the BCS-BEC crossover regime of atomic Fermi gases with strong attractive interactions. We will show that the various definitions lead to FS contours which are not identical, but nevertheless agree with each other to a remarkable degree. All of them violate the Luttinger sum-rule (area enclosed equal to fermion density) and we obtain a detailed understanding of this violation: its magnitude is related to the SC gap function and its sign to the topology of the FS.

Fermi surface criteria.—It is perhaps not appreciated that the question of the “underlying FS” in the SC state is nontrivial, because in BCS theory the answer appears to be simple. In the BCS state one can look at \(G^{BCS}(k, z) = (z + \xi_k)/((z^2 - E_k^2)\) and ask where \(G^{BCS}(k, 0) = 0\). The resulting surface coincides with \(\xi_k = 0\), the normal state FS on which the pairing instability takes place. Thus it is tempting to use \(G(k, 0) = 0\) in a more general setting to define the SC state FS. This is analogous to Luttinger’s definition except \(G(k, 0)\) changes sign through a zero in the SC, instead of a pole in the normal case. However, it is important to note [9] that there is no analog of Luttinger’s theorem for SCs. One can write the Luttinger-Ward functional in terms of the Nambu Green’s function matrix \(\hat{G}\), and try to generalize Luttinger’s proof [1]. However, \(\text{Tr}\hat{G}\) only constrains the difference \((n_1 - n_f)\), which is trivially zero in our case [11], and not the sum [12]. This is related to the fact that spin \(S_z^{\text{total}}\) is conserved in the SC state but the number is not. Thus one cannot show in general that the
surface [10] $G(k, 0) = 0$ in the SC state encloses $n$ fermions. We will come back later to why the Luttinger count nevertheless seems to work in BCS theory.

We next turn to various alternative definitions of the FS. (i) ARPES measures [13] the one-particle spectral function $A(k, \omega) = -\text{Im}G(k, \omega + i0^+)/\pi$ and thus one way to define the “underlying FS” is to look at $A(k, \omega = 0)$ to map out the locus of maximum ARPES intensity. We also describe below a closely related minimum gap locus, also motivated by ARPES experiments [2,3]. (ii) In the cold-atom experiments [8], it is possible to measure the momentum distribution, and therefore we also discuss the moment (somewhat ad hoc) but well defined) criterion $n(k) = 1/2$ to define a surface that separates states of high and low occupation probabilities. (iii) We show below that the quasiparticle excitation spectrum, even in the strongly correlated SC state, is given by $E_k = (\xi_k^2 + |\Delta_k|^2)^{1/2}$ where $\xi_k$ is the renormalized dispersion and $\Delta_k$ the gap function. We then look at the contour defined by $\xi_k = 0$ to define the “FS”. In addition to comparing the contours obtained by various definitions, we also discuss the extent to which these results differ from $G(k, 0) = 0$.

(a) High $T_c$ Superconductors.—We describe the strongly correlated $d$-wave superconducting ground state and low-lying excitations using a variational approach [14,15] to the large $U$ Hubbard model $H = -\sum_{\sigma}n_{\sigma}^a c_{\sigma}^{\dagger} c_{\sigma} + U \sum_{\sigma}n_{\uparrow}n_{\downarrow}$ on a 2D square lattice. We choose [16] the bare dispersion $\xi_k = -\gamma_k + t'\lambda_k$ with $\gamma_k = 2(\cos k_x + \cos k_y)$ and $\lambda_k = 4\cos k_x \cos k_y$. We work at an electron density $n = 1 - x$ with hole doping $x \ll 1$. Our variational ground state is $|\psi_0\rangle = \exp(-i\tilde{S})|\text{BCS}\rangle$, where $|\text{BCS}\rangle$ is the BCS wave function with $d, z, \cdots$ pairing, the projection operator $\mathcal{P} = \prod_i (1 - n_i^{\dagger}n_i)$ eliminates all double-occupation and finite $t'/U$ corrections are built in through $\exp(-i\tilde{S})$ [15]. Here we present the results of a renormalized mean field theory (RMFT) using the Gutzwiller approximation [17] which are in excellent agreement [see Figs. 1(b) and 1(c)] with those obtained using the variational Monte Carlo (MC) [15] method which treats projection exactly. The RMFT approach has advantages over MC calculations for our present investigation since we get much better agreement.

In the RMFT we minimize $\langle \mathcal{H} - \mu N \rangle$ to obtain self-consistency equations for the gap function $\Delta_k = \Delta(\cos k_x - \cos k_y)/2$ and Fock shift $\chi_k$ [18]. These in turn determine the BCS factors $\nu_k^2 = 1 - u_k^2 = (1 - \xi_k/E_k)^2/2$. The renormalized dispersion $\xi'_k = g_i\xi_k - \chi_k - \mu$ incorporates bandwidth suppression by the Gutzwiller factor $g_i = 2x/(1 + x)$, the Fock shift $\chi_k$, and the (Hartree shifted) chemical potential $\mu$. $E_k = (\xi'_k^2 + |\Delta_k|^2)^{1/2}$ is the excitation energy [17] of the projected Bogoliubov quasiparticle (QP) state $|k\rangle = \exp(-i\tilde{S})\mathcal{P}\gamma_{k\uparrow}^\dagger|\text{BCS}\rangle$, where $\gamma_{k\uparrow}^\dagger = (u_k c_{k\uparrow}^\dagger - v_k c_{-k\downarrow})$.

The spectral function is of the form [19] $A(k, \omega) = A^{\text{coh}}(k, \omega) + A^{\text{inc}}(k, \omega)$. The coherent part $A^{\text{coh}}(k, \omega) = Z(k)[\nu_k^2 \delta(\omega - E_k) + \nu_k^2 \delta(\omega - E_k)]$ where the quasiparticle residue $Z(k) = g_i - \frac{i}{2}\sum_k u_k v_k^* \nu_k^2 - \frac{1}{2}\nu_k^2 \sum_k u_k v_k^*$ with $g_i = 2x/(1 + x)$ and $\nu_k^2 = 4x/(1 + x)^2$. The coherent weight $Z$ decreases monotonically with underdoping, vanishing as $x \to 0$, as seen in Fig. 1(c), while the incoherent spectral weight $A^{\text{inc}}$ increases with decreasing $x$ as required by rigorous sum rules [19].

Renormalized dispersion.—The form of the excitation gap $E_k = (\xi'_k^2 + |\Delta_k|^2)^{1/2}$ suggests that we identify $\xi'_k = 0$ as the “underlying FS.” We must emphasize that despite the BCS-like form of $E_k$, theory neither assumes nor implies the existence of sharp quasiparticles in the normal state with a dispersion $\xi_k$. In Fig. 2 we plot the $\xi_k = 0$ contours for various $x$ along with others which will be discussed below. We see that the $\xi_k = 0$ contours are holelike—closed around $(\pi, \pi)$—for small $x$, but electronlike—closed around $(0, 0)$—for $x > 0.16$. The precise $x$ at which the FS topology changes is a sensitive function of the bare parameter [16] $t'/t$.

Minimum gap locus.—We plot in Fig. 3 $A(k, 0)$ which is the zero-energy ARPES intensity. We can neglect the incoherent weight at $\omega = 0$ and write $A(k, 0) \approx Z(k)[\Gamma(\xi_k^2 + \Gamma^2)$, where the $\delta$ function is broadened with a small $\Gamma$. We now follow a procedure developed in analyzing ARPES experiments [2]. Various cuts through $k$ space are taken perpendicular to $\xi_k = 0$. On each of these cuts we determine the location of the maximum $A(k, 0)$, which is also the same as minimum gap $\xi_k$ [ignoring the negligible $k$ dependence of $Z(k)$ on this locus]. The locus of $\text{min}E_k$ defines the “minimum gap locus.” We see from Fig. 2 that the curve $\xi_k = 0$ and the “minimum gap locus,” although not identical, are very similar at every doping. In fact this difference is not visible in Fig. 2 as it is less than the width of the lines used.

Momentum distribution.—We calculate $n(k) = \langle c_{k\sigma}^\dagger c_{k\sigma} \rangle$ and find the result $n(k) = Z(k)\nu_k^2 + \tilde{n}(k)$, where first term comes from coherent quasiparticles and the second term $\tilde{n}(k) = (1 - x)^2/(2(1 + x)) + O(t'/U)$ is the incoherent contribution [20]. The evolution of $n_k$ with
contours enclose areas different from includes incoherent spectral weight \[ 22 \]. function, while \( n_L '' \) depend only on the coherent part of the spectral function of hole doping \( x \). The contours plotted are (1) \( \xi_k = 0 \), (2) \( n(k) = 1/2 \), and (3) the noninteracting Fermi surface \( \epsilon_k = \mu^0 \). Note that the “minimum gap locus” (see text) is indistinguishable from \( \xi_k = 0 \). (d) The fractional difference \( \delta n/n \) between the area enclosed by \( \xi_k = 0 \) and \( n \) plotted as a function of hole doping \( x \).

Fig. 3 (color online). ARPES intensity maps at zero energy \( A(k, 0) \) for \( x = 0.15 \) (left) and \( x = 0.35 \) (right).

doping \( x \) is shown in Fig. 1(a). We choose the contour \( n_k = 1/2 \) \[ 21 \] to (somewhat arbitrarily) separate states of high and low occupation and its variation with \( x \) is plotted in Fig. 2. We see that unlike \( \xi_k = 0 \), the \( n(k) = 1/2 \) contour does not exhibit a change in topology and is “electronlike” down to very low doping. This qualitative difference arises because \( \xi_k = 0 \) and the “minimum gap locus” depend only on the coherent part of the spectral function, while \( n(k) \) is an energy-integrated quantity that includes incoherent spectral weight \[ 22 \].

Luttinger count.—We see from Fig. 2 that the various FS contours enclose areas different from \( n \), which is the area enclosed by the noninteracting FS \( \epsilon_k = \mu^0 \). In Fig. 2(d) we plot the difference between the area enclosed by \( \xi_k = 0 \) and \( n \) as a function of hole doping \( x \). We see that, in general, this difference is nonzero when the system exhibits SC long-range order (\( 0 < x < 0.4 \)) \[ 23 \]. For \( x < 0.16 \), the \( \xi_k = 0 \) contour is holelike and we find an area enclosed greater than \( n \), while for \( 0.16 < x < 0.4 \), this contour is electronlike and the enclosed area is less than \( n \). The \( x > 0.4 \) ground state is a normal Fermi liquid and the Luttinger sum-rule is valid \[ 24 \].

A simple way to understand the variations seen in Fig. 2(d), which include both Gutzwiller and Hartree-Fock renormalizations, is not immediately obvious. However, the BCS-BEC crossover analysis below will give us clear insight into both the (small) magnitude of the violation observed here and the relation of its sign to the FS topology.

Zeros of \( G \).—Next we compare the FS contours obtained above with the surface \( G(k, 0) = 0 \), though we note that the latter is not of direct experimental relevance. From the form of \( A(k, \omega) \) obtained above, we get

\[
G(k, 0) = -Z_k \xi_k/E_k^2 - \int_{-\infty}^{+\infty} d\omega A^{inc}(k, \omega)/\omega. \tag{1}
\]

It is clear that \( \xi_k = 0 \) corresponds only to the first term \( G^{\text{out}}(k, 0) = 0 \), and not \[ 22 \] to a zero of the full \( G \). From the sum-rule constraints \[ 19 \] on \( A^{inc} \) one can show that the integral above is necessarily negative. It then follows that the zeros of \( G \) correspond to \( \xi_k > 0 \). This implies that for small \( x \), where \( \xi_k = 0 \) is holelike, \( G = 0 \) gives an even larger violation of Luttinger count.

(b) BCS-BEC crossover.—The evolution of a Fermi gas from a BCS paired superfluid to a BEC of composite bosons has now been realized in the laboratory \[ 5 \] using a Feshbach resonance to tune the s-wave scattering length \( a_s \). The dimensionless coupling \( g = 1/k_f a_s \) can be varied from large negative (BCS limit) to large positive (BEC limit) values, with unitarity (\( g = 0 \)) being the most strongly interacting point in the crossover. We use the \( T = 0 \) crossover theory \[ 6,7 \] to gain further insight into the question of the “underlying Fermi surface.”

The structure of the \( T = 0 \) Green’s function \[ 10 \] in this case is exactly the same as Eq. (1), with an excitation spectrum \( E_k = (\xi_k^2 + \Delta^2)^{1/2} \). In the well-controlled limit of large dimensionality, treated within dynamical mean field theory \[ 25 \], \( Z \) is close to unity and the incoherent spectral weight is small even at unitarity. Thus, to make our point in the simplest possible manner, we work with Leggett’s variational ansatz \[ 6,26 \]. Even at this level, where incoherent weight vanishes, the implications of the various FS definitions are very interesting.

It is then easy to show analytically \[ 27 \] that all the definitions investigated above yield the same surface in \( k \) space for all values of the coupling \( g = 1/k_f a_s \). This is given by the bare dispersion \( \hbar^2 k_f^2/2m = \mu(g) \), where the chemical potential \( \mu \) strongly renormalized from its noninteracting value. Even in the weak-coupling BCS limit, \( \mu \) is less than the noninteracting \( \epsilon_f \) by an exponentially small amount of order \( \Delta^2/\epsilon_f \) and the Luttinger count is violated. This violation becomes increasingly severe with increasing \( g \); as the gap increases, \( n(k) \) broadens and \( \mu \) decreases; see Ref. \[ 26 \]. Eventually on the BEC side of unitarity (\( g \sim 1 \),
μ goes negative and the surface \( \xi_k = 0 \) shrinks to \( k = 0 \), beyond which one enters the Bose regime. To summarize: the “underlying FS” does not enclose the total number density \( n \), its volume decreases monotonically with \( g \) and for \( g \) greater than a critical value it is zero.

This analytical solution is modified quantitatively by correlation effects beyond the Leggett theory, but qualitative effects like the decrease in \( \mu \) and broadening of \( n(k) \) with increasing \( g \) persist, as also seen in both numerical [28] and experimental studies of \( n(k) \) [8].

We now see how the renormalization of the chemical potential in the presence of a SC condensate directly leads to a violation of the Luttinger count. For not too large attraction, the violation has a relative size \((\Delta / \epsilon_f)^2\), and a negative sign for a particlelike FS; i.e., the underlying FS encloses a smaller area than the noninteracting FS. To see how the sign changes for a hololike FS, we look at the BCS-BEC crossover in a lattice model such as the attractive Hubbard Hamiltonian [29]. It is straightforward to show, using a particle-hole transformation, that sign of the \( \mu \) renormalization reverses going from a particlelike to a hololike FS. Thus we find that for a hololike FS, the underlying FS in the SC state encloses a larger area than the bare FS. These are exactly the effects seen in the strongly correlated \( d \)-wave SC in Fig. 2(d).

Conclusions.—We have analyzed various criteria for the “underlying FS” in the \( T = 0 \) SC state. We have shown that a FS deduced from a SC state measurement necessarily violates the Luttinger sum-rule and does not, in general, enclose \( n \) fermions. We have gained detailed insights into the magnitude and sign of the violation. Our results are of most interest for the high-\( T_c \) cuprates, where they can be tested in ARPES experiments, provided one can independently determine the electron density. The existing ARPES data [4,30] on LaSrCuO show a small violation of the Luttinger count with a sign change, consistent with our results. We should emphasize that our theoretical considerations make no statement about the finite temperature non-Fermi-liquid normal states.

We acknowledge useful discussions with J.C. Campuzano, A. Fujimori, and S. Sachdev. We thank P.W. Anderson for sending us a copy of Ref. [31] while we were preparing this Letter. The results of Ref. [31] are very similar to those in part (a) of our Letter.

[10] Here \( G \) is the \( G_{11} \) element of the Nambu matrix \( \mathbf{G} \).
[12] \((n_1 + n_i)\) is related to \( Tr(\mathbf{G} \bar{\tau}_i)\), for which the Luttinger argument [1] fails.
[16] For our numerical results we choose typical values for the parameters making no statement about the finite temperature non-Fermi-liquid normal states.
[18] The full expression will be published separately; R. Sensarma, M. Randeria, and N. Trivedi (unpublished).
[19] On the zone diagonal \( m_0 \) has a jump discontinuity of \( Z(k_n) \) at the nodal \( k_n \); see Fig. 1. The FS crossing on the diagonal is defined by \( k_n \), rather than \( n(k) = 1/2 \).
[20] Note that for a normal Fermi liquid, it follows from the analog of Eq. (1) that incoherent spectral weight has no effect on the Luttinger FS \( G^{-1}(k,0) = 0 \).
[21] The Mott insulator [15] at \( x = 0 \) has \( Z = 0 \) and is outside the scope of discussion here.
[22] In the \( \Delta = 0 \) Fermi-liquid limit, one can show analytically from that all the contours are identical.
[25] From [26] we see \( G(k,0) = -\xi_k/E_k \) and \( n(k) = (1 - \xi_k/E_k) = 1/2 \) are the same as the \( \xi_k = 0 \), as is the minimum gap locus deduced from \( E_k \). Inclusion of the effects of \( A^{nk} \) in (1) will lead to differences between the various surfaces though, as discussed in case (a).