Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Physics and Chemistry of Solids 67 (2006) 1–5
JOURNAL OF PHYSICS AND CHEMISTRY OF SOLIDS
www.elsevier.com/locate/jpcs

Theory of asymmetric tunneling in the cuprate superconductors

P.W. Anderson *, N.P. Ong

Department of Physics, Princeton University, Princeton, NJ 08544, USA

Abstract

We explain quantitatively, within the Gutzwiller-Resonating Valence Bond theory, the puzzling observation of tunneling conductivity between a metallic point and a cuprate high-T_c superconductor which is markedly asymmetric between positive and negative voltage biases. The asymmetric part does not have a 'coherence peak' but does show structure due to the gap. The fit to data is satisfactory within the over-simplifications of the theory; in particular, it explains the marked 'peak-dip-hump' structure observed on the hole side and a number of other qualitative observations. This asymmetry is strong evidence for the projective nature of the ground state and hence for '*t-J*' physics.

In the conventional, BCS superconductors, the most complete and convincing evidence for the phonon mechanism came from the tunneling spectrum—the tunneling conductivity as a function of junction voltage. The features in this spectrum were shown to be uniquely caused by the anomalous self-energy (the 'gap') and gave unequivocal evidence for its origin in the exchange of phonons. It has been a disappointment that tunneling spectroscopy has so far given us no such evidence for the cuprate superconductors.

One of the puzzling experimental features about the tunneling in cuprates is the fact that the tunneling conductivity is markedly asymmetric as a function of voltage. This was observed even in early crude attempts but became serious when vacuum tunneling from STM points achieved clean results in good detailed agreement (except for this fact) with the expected d-wave density of states [1–3]. It is particularly interesting when one realizes that asymmetries are rare to non-existent in most metal-to-metal contacts, and are predicted not to exist (except for slow, continuous variations of tunneling probabilities) within Fermi liquid theory. The Gutzwiller-projected mean-field-theory [4,5] is not a Fermi liquid theory and can—and does—exhibit asymmetry. Hirsch [6] has drawn attention to the phenomenon in cuprates and proposed that the asymmetry arises from an intrinsic energy dependence ('slope') of the gap. Our theory follows from entirely different assumptions.

The rarity of other examples of structure in tunneling—other than the well-known Giaevar effect of the appearance of the superconducting gap—is a consequence of two remarkable theorems. The first was proved by Harrison [7] essentially in order to explain why Giaevar did not see band structure effects in normal metals. Harrison showed that the tunneling probability between two states, k in metal A and k' in metal B, evaluated in WKB approximation, contains a factor $v(k)v(k')$ from the 'attempt frequency'—where $v(k)$ is the velocity—in addition to the WKB integral. (The theorem is actually more general than WKB, but this will do). This factor cancels against the density of states, which is proportional to $1/v$. Although there may be prominent density-of-states anomalies in the band structure near to the Fermi level caused by a Van Hove singularity, they will not show up strongly in tunneling.

The second theorem is due to Schrieffer [8]. This is particularly useful in the BCS case, but has some bearing in the present one. Schrieffer pointed out that in most situations the tunneling probability is spread over a wide range of k-values, so that one must integrate the single-particle Green's function that appears in the tunneling conductivity over the variable k. This may be converted into a contour integral around the pole in the Green's function at the quasiparticle energy, and simplifies the tunneling density-of-states in the BCS case to $N(E)=N(0)\mathrm{Re}\left[E/\sqrt{E^2-\Delta(E)^2}\right]$ where $\Delta(E)$ is the gap function evaluated at the energy of the quasiparticle pole and $E^2=\left[\varepsilon_k+\sum(k,E)\right]^2+\Delta(k,E)^2$.

The result is not quite so clean in cuprates, where self-energies can be assumed to be k- as well as E-dependent, but this modification seems only to require a factor of $[1+(\partial\Sigma/\partial\varepsilon_k]^{-1}$, which is likely to vary rather smoothly, at least near the coherence peaks.

In the Gutzwiller mean-field-theory [5], we start from the approximation that the ground state of the *t-J* model Hamiltonian is a projected BCS product function, chosen by

* Corresponding author. Tel.: +1 609 258 5850; fax: +1 609 258 1006.
E-mail address: pwa@pupgg.princeton.edu (P.W. Anderson).

0022-3697/$ - see front matter
doi:10.1016/j.jpcs.2005.10.132

minimising the energy over all such functions. The *t-J* Hamiltonian is arrived at by a canonical transformation [9] of the Hubbard Hamiltonian, viz.

$$H_{tJ} = e^{iS} H e^{-iS} = \hat{P} T \hat{P} + J \sum_{<i,j>} S_i \cdot S_j, \quad (1)$$

where T is the kinetic energy and $\hat{P}$ the Gutzwiller projection operator

$$\hat{P} = \prod_i (1 - n_{i\uparrow} n_{i\downarrow}). \quad (2)$$

The exchange term is not projected because it automatically remains within the subspace defined by $\hat{P}$, and hence commutes with it. The eigenstates of the *t-J* Hamiltonian are necessarily projected one-electron functions, so it is natural to approximate them with product functions. Thus, the ground state variational function is

$$|f\rangle = e^{iS} \hat{P} |\Phi\rangle \quad (3)$$

where $\Phi(\Delta, \mu)$ is the BCS product function

$$|\Phi\rangle = \prod_k (u_k + v_k c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow} |0\rangle, \quad (4)$$

and u_k and v_k are the variational parameters, determined by an effective BCS Hamiltonian that gives us a gap equation as discussed in Ref. [4]. This gap equation is actually the equation for the excitation energies of Gutzwiller-projected excited-state wave functions

$$|\Phi_{k\sigma}\rangle = e^{iS} \hat{P} \gamma^\dagger_{k\sigma} |\Phi\rangle, \quad (5)$$

where

$$\gamma^\dagger_{k\uparrow} = u_k c^\dagger_{k\uparrow} - S^\dagger v_k c_{-k\downarrow} \quad (6)$$

and the operator $S^\dagger$ creates a ground-state pair. The procedure is entirely analogous to the Hartree-Fock-BCS theory where the ground state is determined by the criterion that all single-Fermion excitations have positive energy. Within this theory, the excitations in Eq. (5) are the low energy single-Fermion excitations, by Koopman's theorem, which may be demonstrated in this case by the same method as in normal Hartree-Fock.

We note that the theory so far, and its manipulations, are only correct because the Hamiltonian conserves particle number, so that we do not need to consider coherence between states with different particle numbers. In order to specify that the number of electrons is correct we may simply fix average occupancies at the appropriate values, x for the empty state and $(1/2)(1-x)$ for the singly occupied ones of given spin; and then we proceed with Gutzwiller approximation based on those occupancies. It is essential, incidentally, to use a BCS function based on the Fermi level for the right number of electrons. A simple way to see this is that the projection process conserves the total number of electrons; a more rigorous proof can be given by node-counting methods. A slightly subtle point here is that the Fermi level has lost its usual thermodynamic meaning, in a sense.

But if, as in tunneling, we need to add or subtract electrons, we must follow Laughlin [10] and introduce a fugacity factor Z for electron pairs. This is easily computed by noting that the ratio of these two occupancies in the original product function is $(1-x)/(1+x)$, so to get the correct occupancies we must correct the normalization by $(2x/(1+x))^{1/2}$. Hence, for a pair of holes the fugacity factor is

$$Z = \frac{2x}{(1+x)}. \quad (7)$$

The resulting wave function, in the form given by Laughlin [10], is Eq. (4) multiplied by a factor $Z^{-(n\uparrow + n\downarrow)}$. (We choose for perspicuity to express Z as the fugacity of holes rather than electrons; the choice is arbitrary).

This wave function may be rewritten in a form which demonstrates the effect of Z more clearly. In each factor $(u_k + v_k c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow})$ the v_k factor creates a pair of electrons, or conversely the u_k factor can be taken as creating a pair of holes; thus in any component of the wave function which contains u_k, we can insert a corresponding factor Z. Thus, instead of Eq. (4) we could write

$$|\Phi\rangle = \prod_k (Z u_k + v_k c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow}) |0\rangle, \quad (8)$$

and then we have absorbed the fugacity factor into Φ. But the individual factors in Eq. (8) are not normalized. To define the appropriate quasiparticle excitations as in Eq. (6), they must be normalized and thus, finally, the appropriate product function must be written

$$|\Phi''\rangle = \prod_k \frac{(Z u_k + v_k c^\dagger_{k\uparrow} c^\dagger_{-k\downarrow})}{\sqrt{u_k^2 Z^2 + v_k^2}} |0\rangle \quad (9)$$

The individual factors define the product $\gamma_k \gamma_{-k}$, so that the individual gamma contains the normalizing factor $(u^2 + Z^2 v^2)^{1/2}$ Note that the inclusion of the Z factors nicely leads to the Z renormalization of the order parameter, the superfluid density, and the kinetic energy [4]. Z plays a role which can be described as the amplitude of the hole-pair wave function—or at least the relative amplitude of the part of the pair function which is holes as opposed to spins.

We would like to emphasize that the wave function Eq. (9) is completely identical to that used in previous papers and that no results as to quasiparticle energies or ground state averages are at all modified. The quantities u and v define the starting wave function which is to be projected and its populations modified. We can think of the fugacity factor as part of the Gutzwiller projector, if we like; and like the projector, it does not commute with the fermion operators. Thus, when we need to express the excitations in terms of real fermions added to the system—outside the projector—we must use a modified set of u and v. But the real fermions predominantly go in coherently as single quasiparticles, except for the relatively small term for holes coming from the fluctuations of the opposite-spin occupancy.

We want to present the simplest calculation, since the effect in question is a qualitative one. To this end we set $e^{iS} = 1$, which involves errors of the order J/t which vary smoothly with energy—we are neglecting virtual double occupancy, and will therefore overestimate the asymmetry somewhat.

Ignoring e^{iS}, the quasiparticle wave function can be created from the unprojected product function $|\Phi''\rangle$ by either of the operators $\hat{P}c^{\dagger}_{k\uparrow}$ or $\hat{P}c_{-k\downarrow}$. The former creates it with relative amplitude $uZ(Z^2u^2+v^2)^{1/4}$ the latter with relative amplitude $v(Z^2u^2+v^2)^{1/4}$ But what tunnels in from the metal is not a projected quasiparticle but a real one. We can think of the STM point as instantaneously depositing a particle or hole into the Wannier function at the Cu atom under the point, and we then Fourier resolve the Wannier function amplitude at a time $t=0$ into excitations in the superconductor. Thus, the operators which we must consider, acting on our assumed ground state, are $c^{\dagger}_i\hat{P}$ and $c_i\hat{P}$. For present purposes we can discuss only the site operators, later Fourier transforming to get the momentum space ones.

Consider $c^{\dagger}_{i\sigma}\hat{P}$. This may be divided into its two parts belonging to the forbidden and allowed subspaces:

$$c^{\dagger}_{i\sigma}\hat{P} = (1-\hat{P})c^{\dagger}_{i\sigma}\hat{P} + \hat{P}c^{\dagger}_{i\sigma}\hat{P}. \tag{10}$$

The first term contains only excitations with energy larger than the Hubbard U and does not concern us. The second term is finite only x of the time; it requires that the site i contains an electron in the ground state function $\hat{P}|\Phi''\rangle$. Thus, the probability that it creates an excitation in the allowed manifold is multiplied by x. But it will be important to note that because $\hat{P}c^{\dagger}\hat{P}=\hat{P}c^{\dagger}$, obviously, the part of the electron that does not go into the upper Hubbard band creates only single-quasiparticle excitations in this approximation, thus has only a coherent spectrum.

Now consider $c\hat{P}$. This automatically goes into the allowed manifold, but $c\hat{P}|\Phi''\rangle$ is not exactly the same as the quasiparticle function $\hat{P}c|\Phi''\rangle$ because the latter can contain components where $|\Phi''\rangle$ has $n_i=2$ with probability $(1-x^2)/4$, while these do not appear in $c\hat{P}|\Phi''\rangle$. To adjust the normalization, note that

$$c_{\uparrow}\hat{P} = c_{\uparrow}(1-n_{\uparrow}n_{\downarrow}) = \hat{P}(1-n_{\downarrow})c_{\uparrow}, \tag{11}$$

where the site index has been dropped. The average factor reducing the quasiparticle function is

$$c\hat{P} = \hat{P}c(1-<n_{\downarrow}>) = \frac{1}{2}(1+x)\hat{P}c. \tag{12}$$

Thus, the ratio of the normalization factors for the electron vs the hole spectra is $g_t=Z=2x/(1+x)$. But in this case there is an incoherent spectrum, caused by the three-Fermion operator $[n_{\downarrow}-<n_{\downarrow}>]c^{\dagger}$; we do not believe this is a large effect, but it may cause features in the hole spectrum, particularly in the neighborhood of Δ added to the magnetic resonance energy [13].

For each k and spin, except for the rather small term mentioned in the last paragraph, there is only one quasiparticle wave function that appears in this approximation, that obtained by Gutzwiller projecting the suitable BCS product function. Most of the amplitude is 'coherent', a conclusion quite different from that of Wen [11].

At first sight one would think that the ratio of the tunneling conductivity for the sign of voltage V such as to inject a hole—external electrode positive—to that with the opposite sign would be just $2x/(1+x)$. But actually, quasiparticles are not pure electrons or holes but mixtures of the two, and precisely at the gap energy they are equal mixtures, so that at the gap energy the tunneling conductivity for $+V$ and $-V$ should be equal. The relevant tunnel current can flow either in the form of right-moving holes or left-moving electrons, and in the superconductor it is an equal coherent mixture of the two. But it is important to realize that for a given sign of voltage the two states which are coherent have actually the same charge, so that the hole is accompanied by a ground-state pair.

Following Tinkham [12], we write the tunneling Hamiltonian as the sum of the processes

$$u_{\mathbf{k}}\gamma^{\dagger}_{\mathbf{k}0}c_{\mathbf{q}\uparrow}, \qquad v_{\mathbf{k}}c^{\dagger}_{\mathbf{q}\uparrow}\gamma^{\dagger}_{\mathbf{k}1}$$

plus 2 terms conjugate to these which are negligible in the limit $k_BT \ll \Delta_0$. Here $c^{\dagger}_{\mathbf{q}\uparrow}$ creates an electron in the normal metal, $\gamma^{\dagger}_{\mathbf{k}0}$ creates an excitation in the superconductor, and $u^2_{\mathbf{k}}=(1/2)[1+\varepsilon_{\mathbf{k}}/E_{\mathbf{k}}]$ and $v^2_{\mathbf{k}}=(1/2)[1-\varepsilon_{\mathbf{k}}/E_{\mathbf{k}}]$. At positive bias, only the u process which transfers an electron into the superconductor contributes, while at negative bias, the v term dominates (Fig. 1). The latter represents the breakup of a pair with creation of a quasiparticle and the transfer of an electron out of the superconductor. In either process, we must sum over the particle-like ($k>k_{\mathbf{F}}$) and hole-like branches ($k<k_{\mathbf{F}}$). Because $u>(E_{\mathbf{k}})=v<(E_{\mathbf{k}})$ for 2 states of the same $E_{\mathbf{k}}$ in the 2 branches, the sum leads to a second term with the exchange $u\leftrightarrow v$. Thus, the differential conductance dI/dV is symmetric in the exchange of u and v. However, in contrast to the BCS case, it is not symmetric under exchange of holes and electrons. The no-double occupancy constraint strongly suppresses the electron-injection process (Fig. 1(A)).

For electrons, the current is reduced by the projection factor Z relative to that for the holes. Then the amplitude for a given channel is just the effective u for that channel, and the current its square; we get, taking into account the renormalization

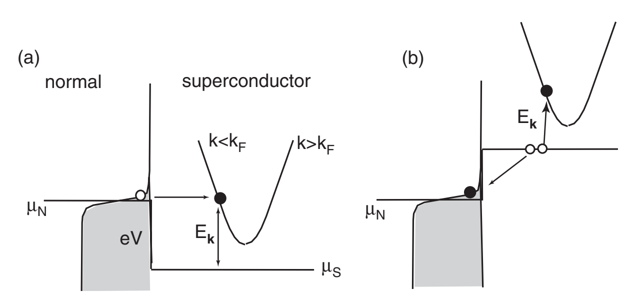

Fig. 1. The tunneling of electrons (panel A) and holes (B) into a superconductor. For positive bias $V>0$ (A), electrons enter the quasiparticle branches $k>k_F$ and $k<k_F$. If $V<0$ (B), the break up of a Cooper pair creates a quasiparticle and ejects an electron into the normal metal (whose occupied states are shown shaded). Strong correlation suppresses the electron tunneling current (A) compared with the hole current (B). At low T, the conjugate processes (same figures with arrows reversed) are negligible.

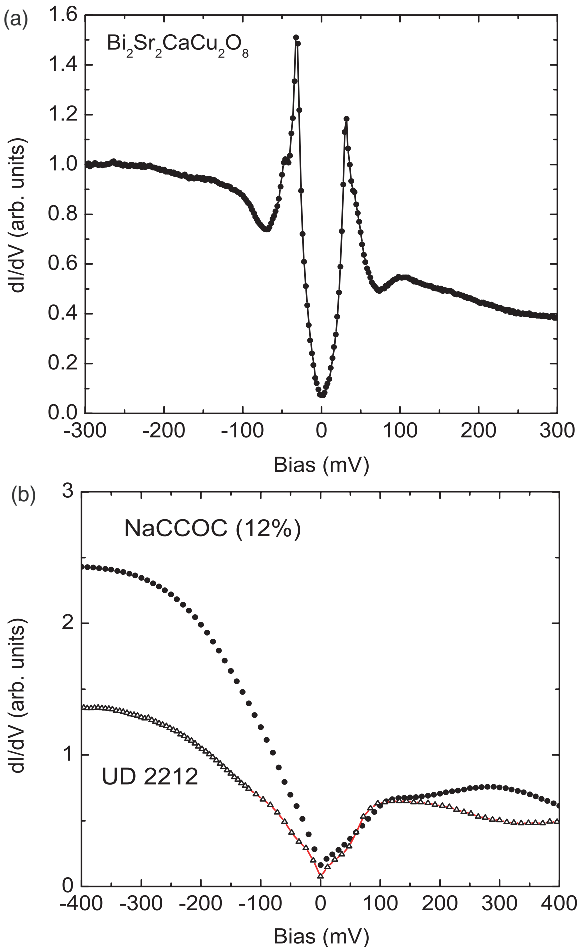

Fig. 2. Tunneling conductance vs voltage for an optimally doped sample of BSCCO (Panel a), and for the oxychloride Mott insulator $Ca_{2-x}Na_xCuO_2Cl_2$ ($x\sim 0.1$) [Panel b]. Curves from the group of Ø. Fischer are quite similar [14]. [Data in Panel a from S H Pan (unpublished), in Panel b from Hanaguri et al. [3]].

factor, that the tunneling current for electrons is (suppressing subscripts k)

$$N_{\mathrm{e}}(E,\Delta)=\frac{\mathrm{d}\varepsilon}{\mathrm{d}E}Z\left(\frac{u^2}{\sqrt{u^2+v^2Z^2}}+\frac{v^2}{\sqrt{v^2+u^2Z^2}}\right). \tag{13}$$

For holes, we do not have the projection factor Z, but the hole amplitude contains the factor Z which can be thought of as the magnitude of the pair wave function. This satisfies the physical requirement that the coherent amplitude for holes and electrons must be the same at least at the same energy, and is also necessary for equilibrium. But the renormalization factor is not identical except at $\varepsilon=0$, $E=\Delta$ and rises as $E\to\infty$ to $1/Z$:

$$N_{\mathrm{h}}(E,\Delta)=\frac{\mathrm{d}\varepsilon}{\mathrm{d}E}Z\left(\frac{v^2}{\sqrt{u^2+v^2Z^2}}+\frac{u^2}{\sqrt{v^2+u^2Z^2}}\right). \tag{14}$$

Eqs. 13 and 14 show that the coherence peaks at $\varepsilon=0$ are identical as predicted, but the ratio of the $E\to\infty$ asymptotes is Z, as expected from simple considerations.

Finally it is necessary to take into account that we have a d-wave gap, which means that the tunnel current must be integrated over the gap distribution

$$\wp(\Delta)\mathrm{d}\Delta=\frac{\mathrm{d}\Delta}{\sqrt{1-\left(\frac{\Delta}{\Delta_0}\right)^2}}, \tag{15}$$

with Δ_0 the gap amplitude. The differential conductance for injection of electrons or holes is then

$$G(E)_{\mathrm{e,h}}=\int_0^{\Delta_0}\mathrm{d}\Delta N_{\mathrm{e,h}}(E,\Delta)\,\wp(\Delta)\mathrm{d}\Delta, \tag{16}$$

with $N_{\mathrm{e,h}}$ given by Eqs. 13 and 14, respectively. For numerical convergence, we add a weak imaginary part to the gap to simulate broadening, i.e. $\Delta\to(1+\mathrm{i}\eta)\Delta$ with $\eta=2\times10^{-3}$.

We have approximated the Fermi Surface by a circle and normalized the maximum gap to 1. The result for the symmetric conductivity in the BCS case has often been displayed, and involves an elliptic function in its analytic form; there is a logarithmic peak at Δ_0, and a linear slope at E near zero. These shapes will appear in the region of the coherence peak and below in the Gutzwiller case, since the asymmetry in Eqs. 13 and 14 is of greater than linear order in $u^2-v^2=\varepsilon/E$. But the renormalizations become appreciable even near $E=\Delta_0$, and especially on the hole side can rise to dominate the peak for the under-doped case. The asymmetry has a pronounced upward cusp at $E=\Delta_0$. A curve from Pan is shown in Fig. 2(a), which refers to a $Bi_2Sr_2CaCu_2O_8$ (BSCCO) sample at optimal doping. It should be noted that it is extremely difficult experimentally to make the conditions exactly identical for $+$ and $-V$, and it is therefore likely that the very small asymmetry

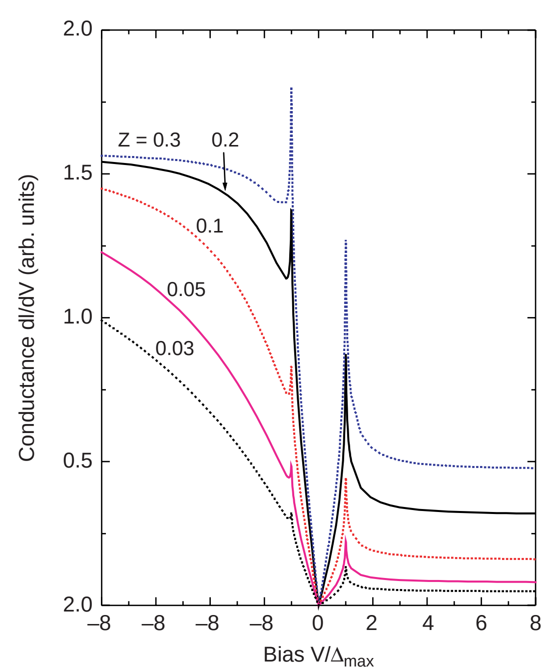

Fig. 3. Computed curves for a sequence of Z's to demonstrate the variation with doping and to allow qualitative extrapolation to $Z=0$. To achieve convergence, a complex gap $\Delta(1+i\eta)$ ($\eta=2\times10^{-3}$) is used in the distribution Eq. (15).

in the coherence peak structure is an experimental artifact, and is actually absent. Ignoring that, the fit to the general course of the asymmetry is remarkable. Other data not shown here confirm the general behavior with doping as well. Fig. 2(b) shows recent results from Davis on the Mott insulator $Ca_{2-x}Na_xCuO_2Cl_2$ ($x\sim 0.1$) which displays the largest asymmetry so far in cuprates.

Another somewhat speculative exercise is to continue Z towards 0, which gives us a conjectured tunneling spectrum for the pure RVB phase, which is our model for the pseudogap state; at least at higher temperatures. In Fig. 3 we show how the variation with doping goes. In the limit as $Z\to 0$, the current (almost all on the hole side) does not extrapolate to an asymptote but continues to rise linearly at high voltage. The ratio of the asymptotes on the two sides is $1/Z$, an observation which seems to accord with most estimates of doping percentages.

The curves which represent regions the experimentalists think are quite underdoped differ from higher dopings in that the symmetrical parts of the curve extend only to rather low voltages, and the coherence peaks are suppressed. The higher-voltage conductivity seems almost completely composed of the asymmetric ‘hump’ behavior and to be dominantly on the hole side. In the same regions, a characteristic ‘4×4’ density wave is observed. In a forthcoming paper we will suggest a mechanism that might relatively suppress the gap antinodes.

We would like to acknowledge extensive discussions with J. C. Davis, and the stimulus of his brilliant analysis of his data. S. H. Pan shared useful unpublished data with us, and Ø. Fischer and L. Greene have also shown us some of their results. Above all, we would like to acknowledge the assistance of Mohit Randeria, Nandini Trivedi and Fu-Chun Zhang in correcting our misconceptions and clarifying our thinking on these issues, without which effort this paper could not have been written.

References

[1] S.H. Pan, et al., Nature 403 (2000) 746.
[2] S. H. Pan, private communication
[3] T. Hanaguri, C. Lupien, Y. Kohsaka, et al., Nature 430 (2004) 1001.
[4] F.C. Zhang, C. Gros, T.M. Rice, H. Shiba, Supercond. Sci. Technol. 1 (1988) 36.
[5] P.W. Anderson, P.A. Lee, T.M. Rice, M.Randeria, Trivedi N., F.C. Zhang, cond mat 0311467: J Phys. Cond. Matt. 16 (2004) R755.
[6] J.E. Hirsch, Phys. Rev. B 59 (1999) 11962.
[7] W. Harrison, Phys. Rev. 123 (1961) 85.
[8] J.R. Schrieffer, Theory of superconductivity, third ed., Perseus Books 1983
[9] C. Gros, R. Joynt, T.M. Rice, Phys. Rev. B 36 (1986) 381.
[10] R. B. Laughlin, et al., Phys. Rev. Lett. 91 (2003) 147003.
[11] W. Rantner, X.G. Wen, Phys. Rev. Lett. 85 (2000) 1722.
[12] M. Tinkham, Phys. Rev. B 6 (1972) 1747.
[13] H.F. Fong, B. Keimer, et al., Phys. Rev. Lett. 75 (1995) 316.
[14] Ch. Renner, Ø. Fischer, Phys. Rev. B 51 (1995) 9208.