Physics of the Resonating Valence Bond (Pseudogap) State of the Doped Mott Insulator: Spin-Charge Locking

P. W. Anderson

Department of Physics, 339 Jadwin Hall, Princeton University, Princeton, New Jersey 08544, USA

(Received 1 March 2005; published 3 January 2006)

The properties of the pseudogap phase above T_c of the high-T_c cuprate superconductors are described by showing that the Anderson-Nambu SU(2) spinors of a resonating valence bond spin gap "lock" to those of the electron charge system because of the resulting improvement of kinetic energy. This enormously extends the range of the vortex liquid state in these materials. A heuristic description of the nonlocal electrodynamics of this pseudogap-vortex liquid state is proposed.

DOI: 10.1103/PhysRevLett.96.017001 PACS numbers: 74.20.−z, 74.72.−h

The hole-doped cuprate superconductors have, over the past decade, been shown to exhibit an unusual phase in the region of their phase diagram immediately above the superconducting "dome" in temperature, and especially for underdoped compositions. This region has been dubbed variously the "pseudogap," or "spin gap" region. Figure 1 shows a generalized phase diagram summarizing the phase boundaries between regions of the complex behaviors exhibited by these systems, and Fig. 2 [due to Ong and Wang [1,2]] is an experimental delineation of the regions where this phase exhibits the behaviors I will discuss later, namely, a pronounced vortex Nernst effect and nonlinear diamagnetic susceptibility, in addition to the pseudogap in density of states and magnetic gap known to exist over the wider region below the "pseudogap temperature" T^*.

In a sense, the existence of a phase of some such sort was predicted quite long before it was observed, in that it was postulated, a few months after the discovery of the high T_c cuprates, that an RVB, a quantum liquid of singlet pairs of electrons ("valence bonds") might exist in an $S = 1/2$ Mott insulating system, from which the high T_c superconductor would develop via doping [3]. Baskaran *et al.* [4], and Fukuyama [5], even drew schematic phase diagrams for such a system which remotely resembled Fig. 1. In a remarkable and prescient paper, Kotliar and Liu [6], using slave boson methods, predicted the existence of a pseudogap phase very much resembling that derived by simpler methods in the present work, but of course they could not compare their results with experiments which did not yet exist. The existence of the pseudogap was first seen in infrared data but not recognized [7], and demonstrated by magnetic [8], tunneling [9], thermal [10], ARPES [11], and other experiments over the years. I particularly want to bring out first the data of Ong *et al.* on the giant Nernst effect in Ref. [1], and the nonlinear diamagnetic susceptibility [11] which demonstrate that over a considerable region this phase exhibits a kind of supercurrent in a magnetic field; and second the infrared data of Timusk [12] which show a marked lengthening of the conductivity relaxation time [13].

The theory I present for this state starts from the renormalized mean field theory [14,15] (RMFT) of the superconducting ground state and uses this as an effective Ginzburg-Landau energy controlling the fluctuations of the variational and other parameters of the state. The variational "gap equations" which determine the energies of the excitations and the energy of the state in RMFT are

$$\Delta(k) = g_J J \sum_{k'} \gamma_{k-k'} \frac{\Delta(k')}{2E_{k'}} E_k^2 = \xi_k^2 + \Delta^2(k) \qquad (1)$$

$$\xi_k = g\varepsilon_k + \varsigma_k = g\varepsilon_k + g_J J \sum_{k'} \gamma_{k-k'} \frac{\xi_{k'}}{2E_{k'}}.$$

Here E is the quasiparticle energy, and ε is its band-theoretical kinetic energy referred to the Fermi energy. g is the kinetic energy renormalization factor due to the modified occupancies in the Gutzwiller-projected BCS state on which RMFT is based. $\langle g\rangle \cong 2x/(1+x)$, x the doping, in a reasonably accurate approximation due to Gutzwiller himself; and g_J is the renormalizing factor for J. Note that for $x = g = 0$, the undoped case, these equations describe the spinon excitations of a pure RVB.

Unfortunately for the simplicity of our exposition, the actual state when $x = 0$ is a Mott insulating antiferromagnet, which is marginally lower in energy than the RVB one, at exactly $g = 0$. Nonetheless we treat the problem of the

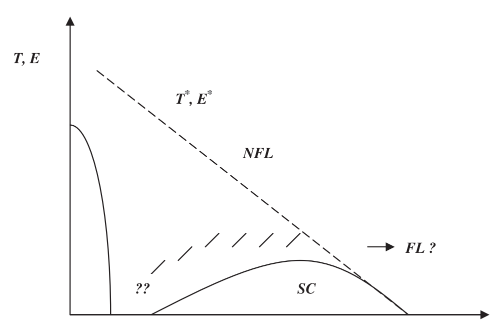

FIG. 1. Generalized phase diagram of the cuprate superconductors.

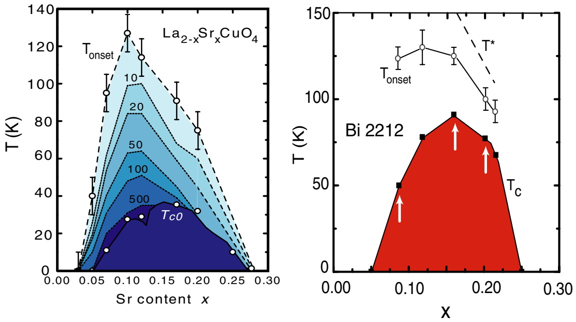

FIG. 2 (color online). The dome and the region of observation of a vortex Nernst effect, as a function of doping for two superconducting systems. [from Refs. [1,11].]

superconducting and pseudogap states using the equations [1] in the approximation $g \cong 0$, since, as we shall show, the kinetic energy gained in the superconducting-RVB state for g finite soon outweighs this small energy advantage. Experimentally, the antiferromagnetism is manifested in this state as the soft "resonance" mode observed in neutron scattering.

The lowest energy translationally invariant state is given, in this limit, by these equations with $\zeta(k) \propto \cos k_x + \cos k_y$ $\varsigma(k) \propto \cos k_x - \cos k_y$ or vice versa. As was emphasized in the early work of Refs. [2,5,14], the solutions, and even the equations, are not unique. It is most evident that Δ and ζ may be interchanged, or in fact any orthogonal rotation between the two is allowed. (It is only in the simple nearest neighbor square lattice without doping that they need to have the same magnitude.) But this is only a symptom of the gigantic symmetry of the solutions, which follows from the fact that the Fermion representation on which they are based is overcomplete. This overcompleteness generates a local SU(2) symmetry [Ref. [3]], which follows from the fact that a spin-up hole and a spin-down electron create the same state of the insulating Heisenberg model in Gutzwiller projection. This symmetry is a true gauge symmetry in that the states generated by it are physically the same state. For translationally invariant singlet BCS states like (1), the gauge symmetry implies symmetry under SU(2) rotations operating on the three Anderson-Nambu pseudospin vectors τ_1, τ_2, τ_3, multiplying, respectively, the real and imaginary anomalous self-energies and the kinetic energies of a gas of "spinons."

But this is only a tiny fraction of the available representations. Also equivalent are states describable as a staggered flux phase, a true π-flux phase, a d-density wave, etc., if one allows translationally noninvariant solutions or T noninvariant ones. What was striking about all good RVB solutions was that in spite of their apparent differences they exhibited four nodes at $\pi/2, \pi/2$ in the spinon energy as a function of momentum, and this is simply because they are the same state. The physics behind this fact is that in order to describe a projected state with no on-site double occupancy each self-energy must average to zero hence have a line node; and that the optimum BCS solution must have as little variation of $|\Delta|^2$ as possible, hence Δ should be the sum of orthogonal functions. This suggests that RVB's will always have nodes in their spectra, even though there is no Goldstone-type argument for nodes in terms of broken symmetry. Thus, in summary, the J term in the t-J model can lead to two and only two self-energies which may multiply two of the three tau vectors; i.e., they are a "dyad" which can be rotated arbitrarily in tau space.

Now we allow g to become finite. Our approach is to set up a second triad of τ's, thinking, if you like, of g as defining the $\tau_3{}'$ direction (that of the charge of the pairs) in this second space. In the RMFT wave function as written in [16]

$$\Psi_g = P \prod_k (g u_k + v_k c*_{k,\sigma}\, c*_{-k,-\sigma})|\text{vac}\rangle, \qquad (2)$$

the phase Φ of the pair wave function is normally taken to be the phase of v/u, but we will ascribe the phase to g and leave u and v real. g is a self-energy in the τ' space whose phase rotates around in the τ_1', τ_2' plane and is conjugate to the charge, thereby defining the third direction. Its meaning at least in the wave function [2] seems to be that it is the hole pair amplitude, which in the ground state is Bose condensed.

We now consider the two triplets of vectors τ and τ' to be oriented at arbitrary Eulerian angles with respect to each other in the abstract space. We may factorize the J term in such a way as to give self-energies along any two, but only two, orthogonal directions in τ' space. The self-consistent solutions for ζ and Δ depend only weakly on these angles for small g, so a good approximation to the effective Hamiltonian for quasiparticles is then

$$h = \sum_k [g\varepsilon_k + (\tau_3{}' \cdot \tau_2)\zeta_k + (\tau_3{}' \cdot \tau_1)\Delta_k](n_k + n_{-k} - 1) + [(\tau_1 \cdot \tau_1{}')\Delta_k + (\tau_1{}' \cdot \tau_2)\zeta_k](c_k{}^* c_{-k}{}^*) + (\text{H.c.}) \qquad (3)$$

Here we derive an effective free energy as a functional of the parameters in the one-electron Hamiltonian which leads to the gap equation—the most important of these parameters being the angles between the two triads of vectors. The "gaps" themselves, Δ and ζ, do not vary much at small g, though Δ, which will become the superconducting gap, will of course decrease appreciably toward optimal doping. The interaction term in the free energy, which will be of the form $\langle|\Delta^2| + |\zeta^2|\rangle/J$, will clearly not change much as a function of angles, so that mostly we have to consider the effective one-quasiparticle energy [3]. The "extended s-wave" solution $\zeta(k)$ to the RVB gap equations strongly resembles the one-particle kinetic energy—which is not surprising since J is a second-order consequence of the kinetic energy. (What is more germane to the special nature of the cuprates is the existence of the second, orthogonal, d-wave solution.) Therefore, it becomes quite obvious that when τ_2 is parallel to $\tau_3{}'$, the kinetic energy is very considerably enhanced, because the coefficient $\xi(k)$, which is the multiplier of the first term in [14], then becomes the sum of two terms of the same structure and the same sign. [This was of course the choice which was made in Ref. [4] and in the original paper [14] on the RMFT.] Therefore, with this choice we maximize the quasiparticle energies, which necessarily minimizes the total energy. The ''locking'' is illustrated in Fig. 3.

This is the central result of the theory: that upon doping the τ_2 axis of the RVB triad *locks* to the $\tau_3{}'$ axis of the Nambu triad of the charge degrees of freedom of the real electrons, so that the τ_1 axis necessarily lies in the $\tau_1{}'$, $\tau_2{}'$ plane and Δ serves as a (necessarily *real*) anomalous self-energy for the actual electrons, which live on the primed triad. The reason for the name ''spin-charge locking'' is the similarity to the mechanism of ''color-flavor locking'' of Wilczek [17].

A second scale is the energy which maintains the broken true phase symmetry. The phase in the $\tau_1{}'$, $\tau_2{}'$ plane is meaningful only as a relative variable and acquires no stiffness from J—a fact which is equally true for conventional superconductors, for which the stiffness $\rho_s(\nabla\phi_{1-2})^2$ is a property of the normal electrons, independent of gap parameters or interactions. The loss of this stiffness may formally be seen as an ''unlocking'' process also, but is not in principle different from a conventional two-dimensional T_c. There are therefore two transitions and three different phases: the low temperature phase is the true superconductor where the two triads may be thought of as locked together; this undergoes loss of phase coherence at T_c, but the $\tau_3{}'$ and τ_2 directions remain locked together in the spin-charge locked pseudogap phase, so that the fluctuating anomalous self-energy remains, retaining its d symmetry and its nodes along the π, π line; and finally, complete unlocking where there may be a pseudogap but it does not show momentum dependence—in this third phase there is no connection between the spin gap structure and the hole or particle nature of the excitations.

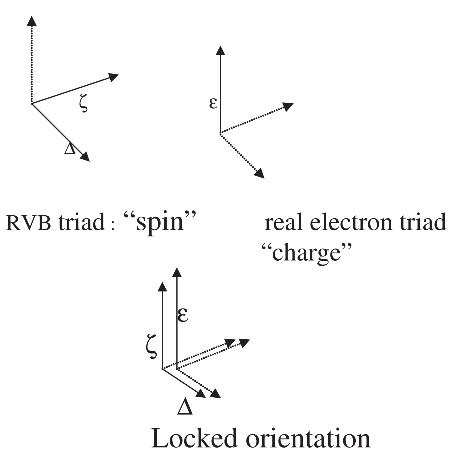

FIG. 3. The charge and spin pseudovector triads, and the relative orientation when ''locked'' (in the superconducting state.)

The transition temperature for the unlocking transition may be estimated from the energy involved. It must lie between T_c, where the phase coherence dies, and $T^* \approx J - xt$, where the RVB gaps appear [18], and we conjecture that it is the crossover or ''onset'' identified by Ong *et al.*. The energy involved is relatively easy to estimate for low doping, where $gt \ll \zeta, \Delta$. Here the average kinetic energy of a state with arbitrary orientation of the τ's is zero, since there will be no correlation between kinetic energy and occupation; while with the locked configuration, the kinetic energy per electron and the locking temperature $T^{*\prime}$ will be of order gt, so that it will rise more steeply than $T_c \approx g\Delta$. This agrees roughly with Ong's observations (see Fig. 2). This rise will stop before $gt \approx \Delta \approx (J - gt)$, where the two contributions ξ to become comparable; this will come at about $1/2$ of the maximum doping. From this point on the calculation of the energy becomes quite complicated and we will postpone it to a later paper. Reference [5] suggests that the onset may not change much with further doping.

The electrodynamics of the partially locked state is in principle the same as that of the vortex liquid state [19], in terms of symmetry. It has in common with that state that there is always an anomalous amplitude, i.e., a ''gap,'' but that the phase of this gap is fluctuating freely; there is a liquid of vortex lines. Not much thought has been given to the responses of such a liquid.

I have concluded that an approximation to its behavior may be obtained in the following way. The instantaneous response to an electromagnetic field is identical to that of an ordinary superconductor, but the current-current correlation decays with a finite correlation time τ. The superconducting response to an electric field is the acceleration equation: $\frac{dJ}{dt} = \rho_s E$; correspondingly, for the locked but nonsuperconducting state we should have

$$J_s = \rho_s \tau E; \qquad \sigma_s = \rho_s \tau. \tag{4}$$

τ, the correlation time for the phase, is a parameter. We have no *a priori* reason to select a magnitude for $1/\tau$, but it certainly should remain less than Δ, otherwise the picture is meaningless; and experimental data on infrared response in the RVB region suggest that the vortices freeze when $1/\tau \approx T = T_c$ [20]. But near T_c the behavior will be critical and complex, and the idea of an overall τ is obviously too simple.

The diamagnetic response is more interesting. Our picture of the pseudogap-vortex liquid state is that the order parameter remains nonvanishing but has no long-range phase order. If one were to introduce a flux quantum instantaneously at a point one would induce supercurrents throughout the sample, including the long-range current $\rho_s(\hbar/m)\nabla\phi \propto 1/r$. Thus the response is fundamentally nonlocal and nonlinear. The energy of the vortex currents varies with B as $B\ln\frac{H_{c2}}{B}$ and one might expect that the Nernst effect thermopower would have a similar variation, and that the susceptibility might also be nonlinear in B and possibly even singular as $B \rightarrow 0$. This nonlinearity was first observed by Ong and co-workers in Ref. [11]; in fact, it was the search for a mechanism for this behavior which led to the present line of thinking. They also observe that the upper critical field H_{c2} is continuous through the melting line and remains large and finite to temperatures where the susceptibility can no longer be measured. This field must be that where the charge becomes unlocked from the RVB, since I cannot see how the field would have much effect on the RVB. I can see no other explanation—for instance, critical behavior—for the remarkable fact that the susceptibility remains nonlinear over such a wide range of temperatures and fields.

It is possible to improve somewhat on the simple expression $B\ln\frac{H_{c2}}{B}$, which is meant only to show the nature of the singularity at $B = 0$. In the d-wave superconductor there is a distribution of gaps and therefore we may approximate its response by summing over a distribution of H_{c2}'s. To get a notion of the form of the variation let us assume that the distribution of Δ is uniform; if we do this we find that the expression for the Nernst thermopower ν becomes

$$\nu \propto B(\sqrt{B} - 1 - \ln[\sqrt{B}]), \quad \text{where } B = H/H_{c2\,\max}. \quad (5)$$

This expression has a family resemblance to some of the data of Ong and Wang.

In summary, I have proposed a theory which accounts for many of the anomalous experimental facts about the pseudogap state, is not falsified by any observations to our knowledge, and connects rather seamlessly to the only successful microscopic theory of the superconducting state. Neither theory places much emphasis on the complications of the various inhomogeneous phases which tend to occur in these systems but which seem to involve smaller energies and weaker perturbations than the more striking and universal effects we discuss. [In fact, one such phenomenon, the "checkerboard," seems to receive a natural explanation within this theory [21].] The crucial step seems to be the idea of visualizing the RVB as a separate entity in the spin Hilbert space which is locked to the electron charge Hilbert space by a relatively weak force in the low doping limit. This makes irrelevant the many "ghost vacua" of the spin system which have led other workers astray. Yet when the spin system chooses a physically different "vacuum"—the Mott insulating antiferromagnetic state, for instance—the whole complex of superconductivity can completely disappear [22].

In terms of the symmetry classification of phases, I have not discovered any "new" entities. The superconducting state is only unconventional in that its quasiparticles are not hole-particle symmetric. The pseudogap phase is "only" an enormous extension of the vortex liquid. Yet each exhibits new and unexpected physical phenomena.

I would like to acknowledge the importance of continued contact with the experimental results and valuable interpretive ideas of N. P. Ong and his group, and also continued discussions with V. Muthukumar. G. Baskaran, G. Kotliar, and P. A. Lee as founders of the gauge theories, and C. Gros, F. C. Zhang, M. Randeria, and N. Trivedi as originators and resuscitators of the RMFT theory, are not adequately acknowledged by the text references.

[1] Z. A. Xu *et al.*, Nature (London) **406**, 486 (2000); Yayu Wang *et al.*, Science **299**, 86 (2003); N. P. Ong *et al.*, Ann. Phys. (Berlin) **13**, 9 (2004).
[2] Yayu Wang *et al.*, cond-mat 0503190; N. P. Ong and Yayu Wang, Physica (Amsterdam) **408C**, 11 (2004).
[3] P. W. Anderson, in *Valence Fluctuations and Heavy Fermions*, edited by L. C. Gupta and S. K. Malik (Plenum, New York, 1987), p. 9; Science **235**, 1196 (1987).
[4] G. Baskaran, Z. Zou, and P. W. Anderson, Solid State Commun. **63**, 973 (1987).
[5] Y. Suzumura, Y. Hasegawa, and H. Fukuyama, J. Phys. Soc. Jpn. **57**, 401 (1988), among many others.
[6] G. Kotliar and J. Liu, Phys. Rev. B **38**, 5142 (1988).
[7] G. A. Thomas *et al.*, Phys. Rev. Lett. **61**, 1313 (1988).
[8] H. Alloul *et al.*, Phys. Rev. Lett. **63**, 1700 (1989).
[9] Ch. Renner *et al.*, Phys. Rev. Lett. **80**, 149 (1998).
[10] J. W. Loram *et al.*, Physica (Amsterdam) **235-240C**, 134 (1994).
[11] D. S. Marshall *et al.*, Phys. Rev. Lett. **76**, 4841 (1996); Hong Ding *et al.*, Nature (London) **382**, 51 (1996).
[12] A. V. Puchkov, D. N. Basov, and T. Timusk, J. Phys. Condens. Matter **8**, 10 049 (1996).
[13] This region may have been first described by J. Corson *et al.*, Nature (London) **398**, 221 (1999).
[14] F. C. Zhang, C. Gros, T. M. Rice, and H. Shiba, J. Supercond. Sci. Tech. **1**, 36 (1988).
[15] P. W. Anderson *et al.*, J. Phys. Condens. Matter **16**, R755 (2004).
[16] P. W. Anderson and N. P. Ong, cond-mat 0405518; Proceedings of the SCE2004 Conference [J. Phys. Chem. Solids (to be published)].
[17] M. Alford, K. Rajagopal, and F. Wilczek, Nucl. Phys. **B537**, 443 (1999).
[18] P. W. Anderson, J. Phys. Chem. Solids **63**, 2145 (2002).
[19] See, for instance, M. Feigelman and L. Joffe, Pizma Zhetf **61**, 71 (1995), and references therein.
[20] C. C. Homes *et al.*, Nature (London) **430**, 539 (2004).
[21] P. W. Anderson, cond-mat 0406038.
[22] I. Bozovic *et al.*, Nature (London) **422**, 873 (2003).