The Relationship Between High-Temperature Superconductivity and the Fractional Quantum Hall Effect

R. B. LAUGHLIN

The case is made that the spin-liquid state of a Mott insulator, hypothesized to exist by Anderson and identified by him as the correct context for discussing hightemperature superconductors, occurs in these materials and exhibits the principles of fractional quantization identified in the fractional quantum Hall effect. The most important of these is that particles carrying a fraction of an elementary quantum number, in this case spin, attract one another by a powerful gauge force, which can lead to a new kind of superconductivity. The temperature scale for the superconductivity is set by an energy gap in the spin-wave spectrum, which is also the fundamental measure of how "liquid" the spins are.

N THE FALL OF 1986, P. W. ANDERSON (1) MADE THE BOLD suggestion that superconductivity in La_2CuO_4 and related materials might be caused by the occurrence in these materials of the "resonating valence bond" state, a hypothetical magnetic liquid state proposed by him (2) in the early 1970s. While neither the resonating valence bond state nor the theory of high-temperature superconductivity it engenders is very well defined at present, I am persuaded that the core of the idea, that the Fermi liquid principle fails in high-temperature superconductors, is right. This has led me to some new perspectives on this subject, which it is the purpose of this article to discuss.

Fermi Liquids and the Resonating Valence Bond

The most appropriate place to begin any discussion of hightemperature superconductivity is the Fermi liquid concept. A Fermi liquid is by definition any system with low-energy excitations similar to those of a noninteracting Fermi sea. It is an empirical fact that all known substances, except perhaps high-temperature superconductors, which are metals in the sense of conducting electricity at zero temperature, are Fermi liquids. For this reason there is a deep-seated belief among solid-state physicists that metals should be Fermi liquids as a matter of principle, even though there is no prima facie theoretical evidence for this. It is almost impossible to demonstrate from first principles that a given material is a Fermi liquid. The equations are simply too complicated. So central is the Fermi liquid concept to our understanding of the solid state that it is implicit in the vocabulary we use and prejudices the questions we ask. The high-temperature superconductivity literature is filled with expressions such as "Fermi surface," "density of states," "Pauli susceptibility," and "electron-phonon interaction," all of which require the context of a Fermi liquid even to make sense.

In light of the overwhelming empirical evidence that all metals are Fermi liquids, it is not surprising that the resonating valence bond concept, although widely respected, is not widely believed. This is unfortunate, for Anderson's reasons for thinking it to be at the bottom of high transition temperature (T_c) superconductivity are compelling (3). Let me paraphrase these as I understand them:

1) High T_c superconductivity occurs in a class of systems, the Mott insulators (4), that we have never understood. It is hard to understand how this could be a coincidence. There is probably some previously unknown property of these systems that causes the effect.

2) The systems in question are inherently magnetic. Stoichiometric La₂CuO₄ is an ordered spin-1/2 antiferromagnet (5) and also an insulator. Doping the material (6) by substituting Sr for about 3% of the La destroys the magnetic order and makes the material a "metal" in the sense of conducting electricity at zero temperature. It is hard to understand how doping at this level could have destroyed all the spins. A more reasonable guess is that the extra holes make ordering more difficult, and that the spins are still present in some sort of "quantum spin liquid" state.

3) The only rotationally invariant spin-1/2 system for which we have an exact solution (7), the linear Heisenberg chain with nearneighbor interactions, possesses a disordered ground state that might well be termed a quantum spin liquid. It is reasonable to adopt this state as a paradigm for the putative spin-liquid state in higher dimension. No one has proved that such states exist, but surely some Hamiltonians can be found in which quantum fluctuations prevent ordering. After all, magnetic ordering is physically similar to crystallization, and helium has both crystalline and fluid phases.

4) The elementary excitations of the Heisenberg chain are known (8) to be neutral spin-1/2 particles possessing a linear energymomentum relation. If the form of these excitations were generic to spin-liquid states, one would expect the charged excitations induced by doping to be very strange, at least within the context of metals as we know them. For example, one possible fate of a hole doped into the material would be to become attached to a neutral spin-1/2 excitation to form a charged spinless particle. Anderson refers to this object, which was invented by Kivelson, Rokhsar, and Sethna (9), as a "holon." He calls the neutral particle a "spinon." Because the holon is spinless, one would guess it to be a boson, in which case it might cause superconductivity by Bose condensing.

The author is in the Department of Physics, Stanford University, Stanford, CA 94305, and is also at Lawrence Livermore National Laboratory, Livermore, CA 94550.

Experimental Properties of High-Temperature Superconductors

There are several significant factors reinforcing the skepticism toward the resonating valence bond idea. One of them is the reluctance of most scientists to abandon thinking that has served them well in the past unless it fails spectacularly. This may or may not have occurred, depending on which experiments one believes. It is regrettably the case that all but a handful of experimental properties of high T_c superconductors can be understood qualitatively in terms of the traditional Fermi liquid theory of superconductors. Let me mention a few of these (10):

1) The spin susceptibility (11) is roughly consistent with Pauli paramagnetism of a Fermi sea containing the number of electrons believed from stoichiometry to have been doped into the material. This carrier density, in turn, is roughly consistent both with the Drude-like conductivity (12) observed in the far infrared and the plasma oscillation induced by it at higher energy. The consistency is only rough because the "band" effective mass of the electrons is not known. Estimates based on experiment range from one to ten electron masses. Attempts to calculate this mass by means of standard band structure techniques are difficult to interpret (13). Stoichiometric La₂CuO₄ comes out to a metal, which it is not.

2) The electrical resistivity (14) above T_c is the size expected of a semiconductor doped to the appropriate level.

3) The transition to superconductivity is associated with a specific heat anomaly (11, 15) related in approximately the right way to the Pauli susceptibility.



Fig. 1. If a spin liquid exists, so should spinons. This is because a system with an even number of spins (**top left**) must be a singlet, while a system with an odd number of spins (**top right**) cannot be. Because the two systems cannot be distinguished when they are large, however, the latter must be viewed as a neutral spin-1/2 excitation of the former. Identification of the half-integral spin as a fractional quantum number suggests that the spin-liquid ground state (**center and bottom right**) is analogous to the fractional quantum Hall state (**center and bottom left**), and that the spinon is analogous to its fractionally charged quasiparticle.

4) Energy gaps have been observed in both the optical absorption (16) and tunneling spectra (17). These gaps are roughly equal and related in roughly the right way to T_c . Both gaps disappear above T_c .

5) The critical field (18) H_c collapses continuously as T_c is approached from below. There is evidence (19) that the tunneling gap also collapses continuously.

6) The zero-temperature London penetration depth λ is roughly consistent (20) with the density of carriers induced by doping. Both λ and the coherence length ξ diverge (18, 21) at T_c .

7) The ac Josephson effect (22) observed in granular samples indicates the presence of a conventional charge-2 order parameter. This is consistent with Cooper pairs but not consistent with Bose condensation of holons. The latter would produce a charge-1 order parameter.

A related and somewhat distressing factor is the inability of the idea to credibly predict anything or, for that matter, even to account for the superconductivity. The most concrete prediction of the approach, namely, that the order parameter should have charge 1, was proved false within a few weeks of its proposition. The situation has become so grim that experimentalists have largely stopped listening to theorists.

The Mott Insulator Problem

By far the most serious impediment, however, is the unwillingness of most solid-state physicists to accept the fundamental intellectual problem a disordered Mott insulator (4) presents. The monoxides of iron, cobalt, and nickel are insulators with similar properties (23). However, cobalt oxide cannot possibly be an ordinary insulator because it has an odd number of electrons per unit cell. All three oxides are, in fact, Mott insulators, materials that insulate solely as a result of Coulomb repulsions between electrons, but only in cobalt oxide is this conclusion inescapable.

Mott insulators are very poorly understood. One frustrating consequence of this is that there are no agreed-upon criteria for identifying them. Thus, even though it is obvious to me that hightemperature superconductors are Mott insulators, it is impossible to convince my disbelieving colleagues that this is the case on purely phenomenological grounds. Lacking an understanding of Mott insulators, we usually assume them to be semiconductors, a wellunderstood class of materials that become metals when doped, and wait for the experiments to tell us otherwise. This would be a perfectly reasonable way to proceed if the experimental results were more clear-cut, for good experiments generally lead to the truth whether ideas motivating them are correct or not. However, it has been the case historically with Mott insulators, and is the case presently with high-temperature superconductors, that the experiments are plagued with materials problems and interpretation ambiguities, so that this strategy does not work well. There is, of course, no reason whatsoever to expect a disordered Mott insulator to be a semiconductor. It is certainly not true that Mott insulators are demonstrably semiconductors (24) described well by a commensurate spin density wave ground state.

Analogy with the Quantum Hall Effect

Let us now ask why the resonating valence bond idea, if it is so insightful, is having so much difficulty accounting for the superconductivity in these materials. The most obvious possibility is that there is a minor error somewhere in its logical development which has led us down a blind alley. Where could it be? The notion of a spin-liquid state seems sound enough. It is an experimental fact that high-temperature superconductors have no magnetic order. It is hard to understand how the spins could simply have vanished. There is certainly no reason to believe that quantum mechanical melting should occur only in one dimension. The most likely source of the problem, therefore, is the identification of the one-dimensional Heisenberg model as an appropriate paradigm.

There are a number of reasons to be suspicious of the Heisenberg chain as a model spin liquid. The most obvious one is that it cannot be ordered, just as a matter of principle, whereas higher dimensional systems can. It is conceivable, for example, that the gaplessness of the spinon spectrum in one dimension simply reflects the system's tendency to be ordered and thus to have a gapless spin wave. Also, it is uniquely the case in one dimension that bosons can be converted to fermions and vice versa by means of canonical transformation (25). Thus, even if the excitations of the higher dimensional spin liquids are analogous to those of the Heisenberg chain, it is not clear what statistics to assign them.

Even while suspecting the paradigm of the Heisenberg chain, one should probably believe in spinons. According to Dzyaloshinskii (26), Landau considered the notion of a liquid state with spin-1/2 excitations so obvious that he did not believe ordered antiferromagnets existed. While it is not completely clear why he thought this, an obvious possibility is illustrated in Fig. 1. Suppose it is established that the ground state of some Hamiltonian is a nondegenerate spin liquid. Then the ground state must be a singlet when the number of spins is even and a doublet when the number of spins is odd. Since there is no long-range order, however, the two systems must be physically equivalent. Therefore the doublet must actually be a spin-1/2 excitation of the singlet ground state. The existence of spinons is



Fig. 2. Whether or not fractional quantization occurs in the quantum Hall problem is determined by the presence or absence of an energy gap in its collective mode spectrum (**top left**). This mode is a compressional sound wave in the fractional quantum Hall state and presumably a spin wave in the spin-liquid state. As the Hamiltonian is changed to induce crystallization (**top right**) the magnetoroton minimum of this mode softens, leading to a divergent susceptibility at the reciprocal lattice vector **G** of the crystal. Viewed in the reduced zone scheme of the crystal (**bottom left**) this spectrum is quite similar to that of a crystal (**bottom right**) except for the presence of a gap at the zone center and the absence of a gap at the zone edge.

28 OCTOBER 1988

on much sounder footing than the analogy with the Heisenberg chain would suggest.

This brings me to what I believe very strongly to be the key idea missing from Anderson's vision of the resonating valence bond. If a spin-liquid state exists, which it probably does, it is expected to have neutral spin-1/2 excitations. But how could this be? In the absence of interactions between the spins, the elementary excitations of the system consist of the act of flipping a spin from down to up. These obviously have spin 1 and are bosons. How could it be that these "elementary particles" of the problem could combine quantum mechanically to make spin-1/2 fermions? In one dimension this point is moot because bosons and fermions cannot be distinguished. In higher dimension, however, the paradox is real. There are two established precedents (27) for turning bosons into fermions in higher dimension, both of which are improbable in the context of this problem. One involves borrowing the half-integral spin representation from isospin degrees of freedom. The other involves adding topological terms to the boson Lagrangian. The latter amounts to changing the laws of physics in a fundamental way. Thus, given that spinons occur in dimension greater than 1, their existence is properly considered miraculous. It implies that the elementary spin-1 excitations have been split in two, with half the excitation appearing in the sample interior and half at the boundary. There is only one identified case of behavior of this kind in nature of which I am aware: the fractionalization of electric charge that occurs in the fractional quantum Hall effect (28). This behavior is so unusual that I find it hard to understand how there could be two distinct phases of matter exhibiting it. I therefore believe that the fractional quantum Hall state is the only possible correct paradigm for the spin-liquid state.

Properties of Incompressible Quantum Fluids

Let us now explore the possibility that the spin-liquid state and the fractional quantum Hall state are one and the same. Here are the generic characteristics of the state as I see them (29):

A featureless liquid-like ground state that is not degenerate.
 Elementary excitations above this ground state that carry fractional charge. This charge is quantized to a particular value

characteristic of the state. 3) An energy gap for making either fractionally charged excitations or collective modes. The collective mode may be thought of either as a pair of fractionally charged particles or as a density fluctuation. Changes to the system's Hamiltonian that preserve this gap also preserve the fractional charge *exactly*.

4) Long-range gauge forces between the fractionally charged particles. These appear in the fractional quantum Hall literature as fractional statistics of the particles. They are also preserved exactly by changes in the Hamiltonian that preserve the gap.

That these are actually properties of the real fractional quantum Hall state is indisputable. We know such a state exists theoretically because a Hamiltonian has been found (30) for which the variational solution I proposed for the problem is exact. We know that excitations out of the state are fractionally charged because we have wave functions for these excitations that are exact (30) in this limit and can prove that their charge is unaffected by changes in the Hamiltonian that preserve the energy gap (29). We also know that the Hall conductance, which is a spectroscopic measure of this charge (31), is exact to within experimental uncertainty (32). We know that these quasiparticles obey fractional statistics both because it can be deduced from their wave functions (29, 33) and because the Hall conductances of the fractional quantum Hall hierarchical states are the correct values (34) to within experimental uncertainty. Like any liquid, the fractional quantum Hall state is difficult to distinguish from a crystal. In particular, it possesses a collective mode (36) that may be thought of either as a compressional sound wave or as an exciton (37) formed when two charged quasiparticles bind. The dispersion relation of this mode is known to have the general form illustrated in Fig. 2. Its deep minimum, named the "magnetoroton" by Girvin, MacDonald, and Platzman (36), who discovered it, occurs at a wave vector corresponding to the interparticle spacing. This gap is a "measure" of how liquid the state is. If the Hamiltonian is tuned so as to induce crystallization, which we know must occur in certain limits, this gap should collapse continuously (38), as appropriate for a second-order phase transition. This has *not* been proved to be the case, but it is very reasonable, and it has certainly not been contradicted by any experiments (39). Three things happen as the gap gets smaller:

1) The state gets increasingly susceptible at the crystallization wave vector. This susceptibility diverges at the crystallization point.

2) The size of the quasiparticles grows. An apt analogy would be the growth of a Cooper pair that results from diminishing the gap of a Bardeen-Cooper-Schrieffer (BCS) superconductor.

3) The energy cost to make a quasiparticle diminishes. The cost is zero at the transition.

The quasiparticle charge maintains its integrity as these things occur as long as the gap is nonzero (29). Let me emphasize again that this behavior is expected on very general grounds and, in particular, is *completely insensitive to what the Hamiltonian is*. The large susceptibility, which is particularly significant for systems on the verge of crystallizing, must lead to intense inelastic scattering at the magnetoroton wave vector. As illustrated in Fig. 2, this could easily be mistaken for Bragg scattering. Since the Goldstone mode of the crystal and its ground-state degeneracy are really the same thing, the



Fig. 3. (Left) Hartree-Fock densities of states for noninteracting particles obeying ν fractional statistics. The fractions $\nu = n/(n + 1)$ are special cases in which a logarithmically divergent energy gap opens up between the occupied (hatched) and unoccupied (unhatched) states. (**Right**) Illustration of the phase associated with interchanging *n* particles obeying ν fractional statistics. The special fractions are precisely the ones for which integral multiples of particles are bosons.

difference between liquid and crystal is simply the presence or absence of the energy gap.

It is reasonable to expect all of these features of the fractional quantum Hall state to have analogs in the spin-liquid state. Thus, the analog of the fractionally charged quasiparticle is the *spinon*, the analog of the compressional sound wave is an antiferromagnetic *spin wave*, the analog of Wigner crystallization is *antiferromagnetic ordering*, and the analog of the magnetoroton gap is a *magnetic fluctuation gap*.

The most obvious implication of the connection between these two states is that the spin-wave spectrum of the magnet must have a gap, as this is the measure of how "liquid" it is. Above this gap the fluctuations should be indistinguishable from those of an ordered antiferromagnet. Although the energy gap has not yet been seen experimentally, magnetic fluctuations similar to those in the ordered phase have been detected in superconducting samples with magnetic Raman scattering by Lyons *et al.* (40). Similar magnetic fluctuations in insulating samples have been seen with inelastic neutron scattering by Endoh *et al.* (6). Since existing experiments do not have the resolution to see this gap, let us guess that it is comparable in size to the one measured in tunneling, or roughly 30 meV. Raman experiments show the maximum spin-wave energy to be roughly 200 meV.

The possibility that the Anderson resonating valence bond state might constitute another example of fractional quantum Hall behavior was suggested to me by D. H. Lee and J. D. Joannopoulos about a year before high-temperature superconductivity was discovered. V. Kalmeyer and I (35) investigated this idea numerically and succeeded in making a very strong case that it makes sense. I must emphasize that we did *not* prove it to be true. Indeed, no one has conclusively proved that a spin-liquid state even exists in any dimension higher than 1. In light of the present experimental situation, however, it seems a bit silly to worry about this. For reasons I have already stated I find the mere consistency of the idea adequate reason for believing it true.

Fractional Statistics of Holons

The most important consequence of the analogy between the fractional quantum Hall and resonating valence bond states is the prediction of a powerful gauge force between the spinons. In the fractional quantum Hall effect, this force, which causes the quasiparticles to obey fractional statistics (34), is known (29) to be a natural concomitant to the presence of fractional charge. A particle carrying electric charge ve moves dynamically as though it carried with it a solenoid containing magnetic flux vhc/e, no matter what the Hamiltonian is, provided that this Hamiltonian can be adiabatically evolved into the "ideal" Hamiltonian without destroying the energy gap. The only thing affected by a change to the Hamiltonian is the size of the solenoid, or equivalently the size of the quasiparticle. Thus, given that the mechanism for quantum number fractionalization in the magnet is the same as that in the fractional quantum Hall effect, such forces are necessarily an attribute of spinons. Furthermore, the fraction of the statistics must be 1/2 because the "charge" of the spinon (35) is $1/2 \hbar$.

Fractional statistics only make sense in two dimensions. It is not clear to me what would happen in a three-dimensional spin liquid, assuming one exists, but a good guess is that the forces become so strong that they confine.

Let us now consider the charged degrees of freedom (9). Whether or not excitations analogous to the holon exist in the fractional quantum Hall effect is not yet clear. The experimental discovery (41)of the "5/2" state and the subsequent confirmation (42) of its magnetic character do demonstrate that charge fractionalization can occur in systems possessing both electric and magnetic degrees of freedom. The quantization of the Hall conductance shows that one of the excitations is a charge-1/2, spin-0 quasiparticle (43). I believe that a charge-0, spin-1/2 quasiparticle also exists, but this has not been demonstrated experimentally.

The main reason to believe that holons exist, however, is that they make so much sense. In trying to conceptualize a charged excitation of a spin liquid, which we know exists because high T_c superconductors can be doped, one immediately runs into the following problem: In order to place an additional electron (hole) on a site it is first necessary to make sure that the electron (hole) already there has the opposite spin. However, because the electron already there is fluctuating quantum mechanically between the up and down states, this requires that one reach in and stop it from fluctuating. This could be accomplished, for example, by projecting the ground state onto the set of states with a given electron down. However, because this creates a large disturbance in the "vacuum," it should be more favorable energetically to create a spinon. Kalmeyer and I (35) found this to be the case when we tested both states using variational wave functions borrowed from the fractional quantum Hall problem. So let us make a spinon. With the spin of the electron in question thus



Fig. 4. Holes (**top left**) and particles (**top right**) in the fractional statistics gas may be thought of as charged vortices. The velocity field of the vortex (**middle left**) falls off as 1/r at large distances, as appropriate for a quantum of circulation $(1 - \nu)\hbar$. The size of the vortex core is $\ell = [2\pi(1 - \nu)\rho]^{-1/2}$, where ρ is the density of the fluid. The action of the density operator ρ_q on the ground state (**middle right**) may be thought of either as a compressional sound wave or as an exciton (**bottom left**) formed from two vortices of the opposite sign. The separation of the vortex is proportional to the exciton momentum **q** and perpendicular to it. The dispersion relation (**bottom right**) crosses over from linear to logarithmic behavior when the vortex separation becomes comparable to ℓ .

28 OCTOBER 1988



Fig. 5. In the presence of a holon condensate, the "elementary particles" of the theory, spinons and holons, cannot be isolated. The physically observable particles must therefore consist of pairs of them. The three possible pairings may be thought of as the excitations $S_q|0\rangle$, $\rho_q|0\rangle$, and $c_{ks}^{\dagger}|0\rangle$.

defined, it is possible to remove the electron in an unambiguous way, thus creating a holon. The holon is spinless because holding the electron down and then removing it is equivalent to holding it up and then removing it. The particle created in this way obviously obeys fractional statistics because it is constructed from a hole and a spinon. However, one can argue more generally that the absorption of the spin of the hole by the vacuum can only have occurred through the fractionalization of the spin quantum number, and therefore must have given rise to a long-range force.

Superconductivity from Gauge Forces

The fractional statistics obeyed by holons has the capacity to cause superconductivity. Unlike the pairing forces in an ordinary superconductor, which are by most measures weak and which have no effect unless they are sufficiently strong to overcome Coulomb repulsions between electrons, the gauge force to which fractional statistics corresponds is strong and leads to charge-2 superfluidity under very general circumstances.

The conclusion that fractional statistics causes superconductivity is based on a recent theoretical study by me (44) of a gas of holons obeying 1/2 fractional statistics and described by the free-particle Hamiltonian

$$\mathscr{H} = -\sum_{j}^{N} \frac{\hbar^{2}}{2m} \nabla_{j}^{2} + \frac{1}{2} \sum_{j \neq k}^{x} V(r_{j} - r_{k})$$
(1)

where V is a pair potential, nominally a Coulomb repulsion. The statement that an energy eigenstate Ψ of this Hamiltonian obeys v fractional statistics means that it takes the form (45)

$$\Psi(z_1,\ldots,z_N) = \prod_{j < k} \frac{(z_j - z_k)^{1-\nu}}{|z_j - z_k|^{1-\nu}} \Phi(z_1,\ldots,z_N)$$
(2)

where Φ is a Fermi wave function and $z = x + i\gamma$ is the position of a particle expressed as a complex number. When written in terms of Φ , the equations of motion become those of fermions moving in the $x - \gamma$ plane and carrying with them a magnetic solenoid containing $(1 - \nu)hc/e$ of magnetic flux. A Hartree-Fock solution of these equations (44) reduces the problem to a gas of noninteracting fermions moving in a uniform magnetic field of strength

$$B = (1 - \nu) \frac{hc\rho}{e} \tag{3}$$

where ρ denotes the particle density. As illustrated in Fig. 3, such a system possesses an *energy gap* in its fermionic excitation spectrum whenever the particle density is an integral multiple of the quantity hc/eB, which occurs in this case when $(1 - \nu)^{-1}$ is an integer. In the fractional statistics gas, this gap turns out to be logarithmically

ARTICLES 529

divergent with the sample size and thus effectively infinite. The energy grows logarithmically because adding a particle polarizes the surrounding fluid in a *vortex* of circulation $(1 - \nu)\hbar$. The condition that a gap exist is precisely the condition that an integral number of particles be bosons. The Hartree-Fock solution thus indicates that the ground state of particles obeying ν fractional statistics is a superfluid with a charge- $(1 - \nu)^{-1}$ order parameter. This charge is 2, as appropriate for a superconductor, for $\nu = 1/2$ holons.

The broken symmetry characteristic of a superfluid is not manifestly present in the Hartree-Fock solution. This is a well-known pathology of variational solutions, which is remedied (46) by hybridizing longitudinal collective modes into the ground state. This collective mode, which is physically the same as a compressional sound wave, appears formally in the Hartree-Fock description as an *exciton* (37) formed, as illustrated in Fig. 4, from a hole in the occupied Landau level and a hole in the lowest unoccupied one. It is physically similar to the purely magnetic collective mode shown in Fig. 2. It may be thought of as a vortex-antivortex pair separated a distance $(\hbar c/eB)q$ and possessing momentum $\hbar q$. The dispersion relation of this excitation is linear at long wavelengths.

The fractional-statistics gas will cease to be a superfluid when the interparticle repulsion V becomes too strong, for then the bosons must form a Wigner crystal. For Coulomb interactions this is thought to occur (47) at densities for which

$$\frac{me^2}{\hbar^2 (\pi\rho)^{1/2}} \ge 100 \tag{4}$$

It has been noted by Peters and Alder (47) that this expression is very nearly an equality at the minimum doping density (3%) required to make La₂CuO₄ metallic.

Gap Collapse and Confinement

Having identified fractional statistics as a possible cause of superconductivity, one is placed in the awkward position of having explained too much. The pairing of holons as a result of fractional statistics is inevitable provided that the spinon gap is nonzero. Thus, in a universe containing only holons, superfluidity would be lost through thermal fluctuations of the order parameter, the transition would look something like the λ -point of liquid helium, and superconducting tunneling would be impossible. Since this is clearly not the case experimentally, the theory can be correct only if spinons are central to the spectroscopy and thermodynamics of these materials. Let me now argue that this is expected to be the case.

One of the most striking features of the high T_c experimental phenomenology is how well it fits the BCS theory. It is hard not to be comforted by this, for it indicates fairly strongly that much of the physics of high-temperature superconductors is the same as that of ordinary superconductors. A moment's reflection, however, reveals that this does not tell one very much. Most of the important properties of a superconductor, such as the Meissner effect, the Josephson effect, and the relation between H_c and T_c , are direct consequences of the occurrence of spontaneous broken symmetry. Thus, experiments which measure all these things are in some sense the same experiment. Similarly, the fact that the $k_{\rm B}T_{\rm c}$ and the energy gap measured in tunneling or by optical spectroscopy are comparable in size and the fact that H_c collapses in a roughly mean-field way as $T_{\rm c}$ is approached from below merely indicate that the destruction of superconductivity is a gap-closing transition. There is nothing inconsistent in this. Even in ordinary superconductors, thermal fluctuations of the order parameter will destroy superconductivity unless something else destroys it first. It is just an accident of nature that the gaps of ordinary superconductors are so small that they

collapse at a relatively low temperatures. Now, given that the holon liquid has the properties I have ascribed to it, the only way to destroy its superfluidity, other than by the thermal fluctuation mechanism, is by destroying the fractional statistics. This, however, can only occur if one destroys the *spinon gap*. Thus the question we need to ask is whether raising the temperature destroys this gap.

Before addressing this question it is necessary to consider the physical meaning of the spinon gap in the presence of a holon fluid. An isolated spinon is expected to induce a vortex in the condensate exactly the way an isolated holon does, because a spinon is simply a holon with an electron added to its center. This means that the energy to make an isolated spinon diverges logarithmically with the sample size, and this means that making an isolated spinon is impossible. One therefore has the strange situation, similar to that occurring in baryons, in which the fractionally charged particles of the theory, the holons and spinons, cannot be isolated. The freely propagating, and thus spectroscopically significant, excitations of the system consist of *pairs* of them. As illustrated in Fig. 5, there should be three of these:

1) A spin wave, consisting of two spinons. This is the excitation illustrated in Fig. 2.

2) A charged current, consisting of two holons. My present understanding is that this should be analogous to the action of the density operator on an ordinary superconductor, which is a longitudinal excitation that loses its identity by hybridizing strongly with the Goldstone mode.

3) An electron, consisting of a holon "hole" and a spinon. This is the excitation created in a tunneling experiment.

Each of these particles should be characterized at low energy by a spin and total momentum, as appropriate for a tightly bound state of two particles. Because the impossibility of isolating spinons is a property of the holon fluid and not of the underlying spin system, it does not invalidate the concept of a spinon gap. It merely requires that one detect collapse of the spinon gap through collapse of the *spin-wave gap* which is expected to occur simultaneously.



Fig. 6. Hypothesized behavior of magnetic excitation $S_q|0\rangle$ in a real high T_c superconductor. Note the similarity to Fig. 2. The energy at Γ is comparable to the optical magnon of the antiferromagnetically ordered state. The gap Δ_s at Γ is comparable to the superconducting gap $2\Delta_e$.

SCIENCE, VOL. 242

Let us now consider the question of gap collapse. It is known that disorder, and thus presumably thermally excited collective modes, can continuously collapse the magnetoroton gap in the fractional quantum Hall state to zero. This is known both from theoretical considerations (38, 48) and from the experimental observation (39, 49) that increasing disorder lowers the activation energy for σ_{xx} continuously to zero and then destroys the fractional quantum Hall effect entirely. Because of the nature of self-consistent gap collapse, namely, that thermal excitation of particles across the gap lowers the gap, which allows even more particles to be excited, which lowers the gap still further, and so forth, one can say without knowing any details that the temperature of gap collapse must be comparable to the gap itself. Thus, for the superconductors we would have

$$\frac{\Delta_{\rm s}}{k_{\rm B}T_{\rm c}} \approx 1 \tag{5}$$

where Δ_s denotes the spin-wave gap. The value of 30 meV that I estimated for Δ_s gives a value for this ratio of 3, which is quite reasonable. It is therefore the case that self-consistent gap collapse of the type necessary to destroy superfluidity is expected at a temperature that is about right.

Consequences of Fractional Statistics Pairing

In order to proceed further it will be necessary for me to make some educated guesses about the precise nature of the excitation spectra in a system of this kind. These "gedanken" calculations are a poor substitute for real ones, particularly because most of the significant questions about these materials are quantitative, but it is the best any of us can do at present.

Undoped La₂CuO₄ is known (5) to be antiferromagnetically ordered along the [110] direction in the Cu-O planes. Accordingly, the spin-wave spectrum of the magnetic liquid state should look something like Fig. 6, with the gap Δ_s occurring at the **M** point in the Brillouin zone. This gap is physically analogous to the magnetoroton minimum shown in Fig. 2. It is shown with a parabolic shape because this would be the outcome of a variational calculation based



Fig. 7. Effect of quantum fluctuations on the spin-wave dispersion (**left**) and density of states (**right**). In the absence of fluctuations (dashed) the dispersion relation is quadratic and the density of states constant. In the presence of fluctuations (solid) both of these become linearized. Superconducting fluctuations have the same effect on the electronic density of states.

on the single-mode approximation. However, it is well known (50) that quantum fluctuations tend to linearize a mean-field dispersion relation of this kind in the limit that the gap is small. Let us therefore guess that the dispersion relation near the minimum is roughly of the form

$$p(\mathbf{q}) \cong (\nu^2 |\mathbf{q} - \mathbf{M}|^2 + \Delta_s^2)^{1/2}$$
 (6)

where v is an asymptotic spin-wave velocity. This dispersion relation and the density of states to which it corresponds are shown in Fig. 7.

Let us now consider the behavior of the "electron." Whatever this excitation is, it should have a significant projection onto the state $c_{\mathbf{k}s}^{\dagger}|0\rangle$, where $|0\rangle$ denotes the true ground state of the system. Thus, to calculate its properties, one could either compute the time-ordered Green's function $G_{\mathbf{k}s}(\tau) = -i\langle 0|T\{c_{\mathbf{k}s}(\tau)c_{\mathbf{k}s}^{\dagger}(0)\}|0\rangle$ or use the state $c_{\mathbf{k}s}^{\dagger}|0\rangle$ as a variational ansatz. Unless "electrons" and "holes" interact very anomalously, the Green's functions for spin and density fluctuations, namely, $-i\langle 0|T\{S_{\mathbf{k}}^{\mu}(\tau)S_{\mathbf{k}}^{\nu}(0)\}|0\rangle$ and $-i\langle 0|T\{\rho_{\mathbf{k}}(\tau)\rho_{\mathbf{k}}(0)\}|0\rangle$, with

$$S_{\mathbf{q}} = \frac{\hbar}{2} \sum_{\mathbf{k}} \sum_{ss'} \langle s | \boldsymbol{\sigma} | s' \rangle c_{\mathbf{k}+\mathbf{q},s}^{\dagger} c_{\mathbf{k},s'}$$
(7)

$$\mathbf{p}_{\mathbf{q}} = \sum_{\mathbf{k},s} c_{\mathbf{k}+\mathbf{q},s}^{\dagger} c_{\mathbf{k},s} \tag{8}$$

must have large spectral weight at the energies of free electron hole pairs, as they do in ordinary superconductors. Thus the presence of a soft spin wave at **M** implies that there are also soft electron hole pair excitations with this momentum. Let us therefore guess that the electron spectrum looks something like that depicted in Fig. 8, with a small gap $2\Delta_e$ at the **X** point. This gap should be slightly greater than Δ_s or equal to it, according to whether the electron hole interaction is attractive or repulsive. The fact that this gap is direct implies that it would produce a strong signal in optical reflectivity, as is experimentally the case. It should also be the gap observed in tunneling.

ĥ

I have assumed the electron spectrum to have its minimum at a *point* in the Brillouin zone because this is the most likely outcome of a variational estimate based on the wave function $a_{ks}^{\dagger}|0\rangle$ or a projected version of it. This behavior, which is precluded by the Fermi liquid hypothesis, has also been suggested by Anderson (3) and by Kotliar (51), although on different grounds. It has the capacity to account for the linear-tunneling density of states (52) seen above the gap in most high T_c superconductors. The quadratic minimum produced by the variational calculation would be expected to linearize, much the way the spin-wave dispersion linearizes, as a result of superconducting fluctuations, producing an electron density of states very similar to the spin-wave density of states shown in Fig. 7.

Let us finally consider the electromagnetic response. The soft charged excitation at the **M** point should become a plasmon and thus effectively become invisible. A soft transverse excitation should, however, be visible at **M** as it is at Γ . There should not be a gap in the density fluctuation spectrum at q = 0. In a BCS superconductor, this can be shown to follow from "backflow" corrections (53) to the bare response function, which does have a gap, but it is actually a general consequence of broken symmetry. The measured density fluctuation spectrum must have a strong pole at the Goldstone mode.

An extremely important consequence of the "confining" effect of the holon gas is that it can account for the stabilization of liquid state by doping. One way to understand the tendency to order, which occurs in the undoped limit, is that the magnetoroton

and

minimum depicted in Fig. 2 is negative. As illustrated in Fig. 8, however, this excitation may also be viewed as a pair of spinons bound together. In the presence of the holon gas, the energy of this excitation, and thus Δ_s , will be raised. If the separation of the spinons associated with this excitation is assumed to be $\rho G/2$, where ρ is the holon density and G is a reciprocal lattice vector, then the vortices in the holon fluid are farther apart than their core size, and we may write, up to an unimportant logarithmic term,

$$\Delta_{\rm s} \cong \Delta_{\rm s}^0 + \alpha \, \frac{\hbar^2}{m} \rho \tag{9}$$

where Δ_s^0 is the (possibly negative) value of the gap at zero doping and α is a coefficient of order unity. This is precisely the type of doping dependence of the gap proposed by Anderson (3) and found phenomenologically by Uemura et al. (54). Assuming a value of 10^{-2} Å^{-2} for ρ and a bare electron mass, one obtains 40 meV for this increase, which is of the correct order.

I would like finally to make a remark about gaplessness. It is a fact that the energy gap or gaps of high-temperature superconductors are extremely difficult to measure. It is commonly the case, for example, that a tunnel junction (17) exhibits a smooth transfer characteristic with a small smooth bump where the gap ought to be. As a result, tunneling spectroscopists cannot agree on the value of the tunneling gap to within a factor of 2. Some deny that there even is a gap. Similarly, the value of the gap determined from infrared absorption (16) varies from sample to sample and does not agree very well with the value determined from tunneling. It is not yet clear why these difficulties occur. However, since gapless superconductors are known to exist and to be caused by magnetic impurities, and since defects in the structure of materials with such low carrier densities are bound to be magnetic, it is quite conceivable that hightemperature superconductors are chronically gapless. In light of this possibility, it is important to make clear that gaplessness does not invalidate the fractional statistics concept. Fractional quantum Hall systems are always dirty and thus always in some sense "gapless."



Fig. 8. Hypothesized behavior of "electronic" excitations $c_{ks}^{\dagger}|0\rangle$ and $c_{ks}|0\rangle$. The gap Δ_e is the one measured in tunneling. The energy at Γ is comparable to the Fermi energy of a gas of electrons at the holon density. Because the gap is direct, it may be observed optically.

Nevertheless we know experimentally that the quantum of Hall conductance, and thus the quasiparticle charge, is exactly quantized so long as the effect is not destroyed entirely.

Conclusion

The purpose of considering the experimental implications of the equivalence of the fractional quantum Hall effect and high-temperature superconductivity in this cursory way is not to prove it correct but rather to show that it is not obviously wrong. I am in agreement with Anderson that the mathematical tools required to accurately calculate properties of this state probably do not yet exist. Before making the effort to invent them it is obviously a good idea to find out if the approach makes sense. The existence of a spin-liquid state and the occurrence of charge fractionalization in such a state are, in my opinion, on firm ground. The ability of fractional statistics or its three-dimensional analog to cause superconductivity is less clear-cut, but probably right. Whether or not such things occur in real hightemperature superconductors is problematical. I am persuaded that they do, but this remains to be demonstrated.

REFERENCES AND NOTES

- 1. P. W. Anderson, Science 235, 1196 (1987)
- ., Mater. Res. Bull. **8**, 153 (1973).
- 3 in Frontiers and Borderlines in Many Particle Physics, J. R. Schrieffer and R. A.
- ______, in Frontiers and Borderlines in Many Particle Physics, J. K. Schrieffer and K. A. Broglia, Eds. (North-Holland, Amsterdam, in press).
 N. F. Mott, Proc. R. Soc. London Ser. A 62, 414 (1949); B. H. Brandow, Adv. Phys. 26, 651 (1977); I. Yu. Pomeranchuk, J. Phys. (Moscow) 4, 356 (1941).
 D. Vaknin et al., Phys. Rev. Lett. 58, 2802 (1987); Y. J. Uemura et al., ibid. 59, 1045 (1987); J. M. Tranquada et al., ibid. 60, 156 (1988).
 G. Shirane et al., ibid. 59, 1613 (1987); Y. Endoh et al., Phys. Rev. B. 37, 7443 (1987); Y. Endoh et al., Phys. Rev. B. 37, 7443
- (1988)
- H. A. Bethe, Z. Phys. 71, 205 (1931).
 L. D. Fadeev and L. A. Takhtajan, Phys. Lett. 85A, 375 (1981).
 S. A. Kivelson, D. S. Rokhsar, J. P. Sethna, Phys. Rev. B. 35, 8865 (1987).
- 10. A good review of the experimental literature may be found in T. H. Geballe and J.
- . Hulm, Science 239, 367 (1988). 11. B. Batlogg, A. P. Ramirez, R. J. Cava, R. B. van Dover, E. A. Rietman, Phys. Rev.
- 12
- B 35, 5340 (1987).
 J. Orenstein et al., ibid. 36, 729 (1987); W. Kamaras et al., ibid., p. 733.
 L. F. Mattheiss, Phys. Rev. Lett. 58, 1028 (1987); J. Yu, A. J. Freeman, J.-H. Xu, 13. ibid., p. 1035.
- 14. M. Gurvitch and A. T. Fiory, ibid. 59, 1337 (1987).
- 15. S. E. Inderhees et al., ibid. 60, 1178 (1988).
- Z. Schlessinger, R. T. Collins, D. L. Kaiser, F. Holtzberg, *ibid.* 59, 1958 (1987).
 M. Lee, M. Naito, M. R. Beasley, A. Kapitulnik, in preparation. Previous pointunreliable. See M. D. Kirk et al., Phys. Rev. B 35, 8850 (1987).

- unreliable. See M. D. Kirk et al., Phys. Rev. B 35, 8850 (1987).
 18. D. K. Finnemore et al., ibid., p. 5319.
 19. I. Igouchi, S. Narumi, Y. Kasai, A. Sugishita, Physica 148B, 322 (1987).
 20. W. J. Kossler et al., Phys. Rev. B 35, 7133 (1987); G. Aeppli et al., ibid., p. 7129.
 21. B.-D. Oh et al., ibid. 37, 7861 (1988).
 22. J. S. Tsai, Y. Kubo, J. Tabuchi, Phys. Rev. Lett. 58, 1979 (1987); Jpn. J. Appl. Phys. 26, C701 (1987); J. Niemeyer et al., Z. Phys. B 69, 1 (1987); D. Esleve et al., Europhys. Lett. 3, 1237 (1987).
 23. B. Koiller and I. M. Falicov, J. Phys. C 7, 299 (1974).
- 23. B. Koiller and L. M. Falicov, J. Phys. C 7, 299 (1974).
- K. Terakura, A. R. Williams, T. Oguchi, J. Kubler, Phys. Rev. Lett. 52, 1830 24. (1984)
- E. H. Lieb and D. C. Mattis, Mathematical Physics in One Dimension: Exactly Soluble Models of Interacting Particles (Academic Press, New York, 1966).
 I. Dzyaloshinskii, A. Polyakov, P. Wiegmann, Phys. Lett. 127, 112 (1988).
- 27. F. Wilczek and A. Zee, Phys. Rev. Lett. 51, 2250 (1983), and references therein. Addition of a topological term is the basis for much recent Soviet work. See A. M. Połyakov [*Mad. Phys. Lett.* A3, 325 (1988)] and P. B. Wiegmann [*Phys. Rev. Lett.* 60, 821 (1988)].
 R. B. Laughlin, *Phys. Rev. Lett.* 50, 1395 (1983).
- _, in The Quantum Hall Effect, R. E. Prange and S. M. Girvin, Eds. (Springer, 29 New York, 1987), pp. 233-301.
- F. D. M. Haldane, *Phys. Rev. Lett.* 51, 605 (1983).
 B. I. Halperin, *Phys. Rev. B* 25, 2185 (1982); R. B. Laughlin, *ibid.* 23, 5632 (1981).
- G. S. Boebinger, A. M. Chang, H. L. Störmer, D. C. Tsui, *ibid.* 32, 4268 (1985).
 D. Arovas, J. R. Schrieffer, F. Wilczek, *Phys. Rev. Lett.* 53, 722 (1984).
- 34. B. I. Halperin, ibid. 52, 1583 (1984).
- V. Kalmeyer and R. B. Laughlin, *ibid.* **59**, 2095 (1987).
 S. M. Girvin, A. H. MacDonald, P. M. Platzman, *ibid.* **54**, 581 (1985).

- 37. Yu. A. Bychkov, S. V. Iordanskii, G. M. Eliashberg, Pis'ma Zh. Eksp. Toer. Fiz. 33, Kukushkin, V. Timofeev, K. von Klitzing, K. Ploog, Festkörperprobleme (Adv.
 K. Kukushkin, V. Timofeev, K. von Klitzing, K. Ploog, Festkörperprobleme (Adv.
- Solid State Phys.) 28, 21 (1988).
- 40. K. B. Lyons, P. A. Fleury, L. F. Schneemeyer, J. V. Waszczak, Phys. Rev. Lett. 60, 732 (1988).
- R. Willet *et al.*, *ibid.* 59, 1776 (1987).
 J. Eisenstein *et al.*, *ibid.* 61, 997 (1988)
- 43. F. D. M. Haldane and E. H. Rezayi, ibid. 60, 956 (1988).
- R. B. Laughlin, *ibid.*, p. 2677.
 F. Wilczek, *ibid.* 49, 957 (1982); D. P. Arovas, R. Schrieffer, F. Wilczek, A. Zee, Nucl. Phys. B 251, 117 (1985)
- E. Feenberg, *Theory of Quantum Fluids* (Academic Press, New York, 1969), p. 107.
 D. Peters and B. Alder, in *Computer Simulation Studies in Condensed Matter Physics: Recent Developments*, D. P. Landau and H. B. Schütler, Eds. (Springer, Berlin, Heidelberg, in press); D. M. Ceperley and E. L. Pollock, Phys. Rev. Lett. 56, 351

- (1986); E. L. Pollock and D. M. Ceperley, *Phys. Rev. B* 36, 8343 (1987).
 48. S. M. Girvin, A. H. MacDonald, P. M. Platzman, *Phys. Rev. B* 33, 2481 (1986); A. H. MacDonald, K. L. Liu, S. M. Girvin, P. M. Platzman, *ibid.*, p. 4014.
- 49. G. S. Boebinger, A. M. Chang, H. L. Störmer, D. C. Tsui, Phys. Rev. Lett. 55, 1606 (1985).

- C. Kittel, Quantum Theory of Solids (Wiley, New York, 1963), p. 58.
 G. Kotliar, Phys. Rev. B 37, 3664 (1988).
 P. W. Anderson and Z. Zou, Phys. Rev. Lett. 60, 137 (1988).
 J. R. Schrieffer, Theory of Superconductivity (Benjamin, New York, 1983), p. 226.
 Y. J. Uemura et al., Phys. Rev. B 38, 909 (1988).
- 55. I gratefully acknowledge numerous helpful discussions with S. Kivelson, J. Sethna, Kalmeyer, C. Hanna, L. Susskind, A. L. Fetter, P. W. Anderson, F. Wilczek, B. I. Halperin, J. R. Schrieffer, T. H. Geballe, M. R. Beasley, and A. Kapitulnik. This work was supported primarily by the National Science Foundation under grant DMR-85-10062 and by the NSF-MRL program through the Center for Materials Research at Stanford University. Additional support was provided by the U.S. Department of Energy through the Lawrence Livermore National Laboratory under contract W-7405-Eng-48.

How Do Enzymes Work?

Joseph Kraut

The principle of transition-state stabilization asserts that the occurrence of enzymic catalysis is equivalent to saying that an enzyme binds the transition state much more strongly than it binds the ground-state reactants. An outline of the origin and gradual acceptance of this idea is presented, and elementary transition-state theory is reviewed. It is pointed out that a misconception about the theory has led to oversimplification of the accepted expression relating catalysis and binding, and an amended expression is given. Some implications of the transitionstate binding principle are then explored. The amended expression suggests that internal molecular dynamics may also play a role in enzymic catalysis. Although such effects probably do not make a major contribution, their magnitude is completely unknown. Two examples of recent advances due to application of the transition-state binding principle are reviewed, one pertaining to the zinc protease mechanism and the other to the generation of catalytic antibodies.

VEN A CASUAL SURVEY OF THE CURRENT BIOCHEMICAL literature reveals a rising interest in enzymes. This upsurge is due in part to the advent of site-directed mutagenesis methods, which have now been reduced to an almost standardized collection of laboratory procedures (1) whereby the amino acid sequence of a given enzyme molecule (or any other kind of protein molecule) may be altered by deliberately and precisely mutating the cloned gene encoding that molecule. As a tool for investigating structure-function relations, site-directed mutagenesis is made still more powerful by the use of x-ray crystallography to redetermine the three-dimensional structure of the mutated enzyme molecule and thereby define exactly what has been changed. The large amount

The increasingly widespread application of site-directed mutagenesis techniques, together with steady advances in methods for preparing hybrid enzymes, semi-synthetic enzymes, and even totally synthetic enzyme-mimetic compounds, and most recently for the production of catalytically active antibodies (3), has given birth to a burgeoning new discipline with the optimistic name of enzyme engineering.

Reasons for this growing interest are not hard to find. Among them are the practical possibilities of putting engineered enzymes to work in industrial and medical applications. Also, since most drugs act by modifying or blocking the activity of some enzyme or another, a deeper understanding of suitably chosen target enzymes should lead to major advances in rational drug design. But most compelling is our sheer curiosity about these ingenious molecular machines, operating at the boundary where chemistry just becomes biology.

The phenomenal rate accelerations and specificities of enzymes have intrigued investigators ever since the 1830s when enzymic activity was first observed [see page 8 of (4)]. Over the years numerous hypotheses and ad hoc explanations have been advanced to account for enzymic catalysis, many of them tagged with imaginative names by their proponents. Page lists no fewer than 21 hypotheses (5). But only gradually has it come to be accepted that the most profitable way to think about the problem is the one first clearly stated by Pauling some 40 years ago (6). The basic idea, as simple as it is elegant, results from a straightforward combination of two fundamental principles of physical chemistry: absolute reactionrate theory and the thermodynamic cycle. In this view an enzyme is essentially a flexible molecular template, designed by evolution to be precisely complementary to the reactants in their activated transition-state geometry, as distinct from their ground-state geometry. Thus an enzyme strongly binds the transition state, greatly increas-

The author is a professor of chemistry and biochemistry at the University of California, San Diego, La Jolla, CA 92093.