Superconductor-Insulator transition in 2D Disordered Films

Ke Dong

April, 1999

Superconductor-Insulator transition(SIT)[1] is a result of the competition between the attractive interaction which is responsible for superconducting pairing, and the pair-breaking effects of localization and disorder-enhanced Coulomb repulsion. Superconductivity is a manifestation of long-range phase coherence between electron pair states, while electronic localization involves a spatial limitation of the electron wavefunctions, which would preclude such pairing.

1 Introduction: the role of disorder

Dirty Superconductor

Anderson[2] argued that superconductivity is insensitive to perturbations that do not destroy time-reversal invariance. For s-wave superconductors, because the interaction between the two electrons in a Cooper pair is attractive, it will be energetically favorable if these two electrons stay close, therefore their spin must be in the singlet state, and the electrons have opposite spins $\vec{\sigma_1} = -\vec{\sigma_2}$; it can also be shown that the largest binding energy occurs for electrons diametrically opposite to each other around the Fermi surface, in other words $\vec{k_1} = -\vec{k_2}$. As we know, under time-reversal transformation, both momentums and spins change signs, as a result, the wavefunctions describing Cooper pairs should be time-reversal invariant. If we introduce some magnetic perturbations into our system, the pairing mechanism should be weakened. When we increase the density of the magnetic impurities, superconductivity will soon be destroyed. On the other hand, the effect of non-magnetic impurities is not so dramatic. It is true that these impurities can add or subtract electrons from the band, change the density of states, nevertheless, they can only gradually vary the parameters of electrons. The basic assumption Anderson made is that, in weak interaction region, the problem of solving the electron wavefunction is best solved by first diagonizing the scattering interaction between the electrons and the impurities, and then calculating the phonon interaction between electrons. Then all the superconducting properties will follow if we can find an attractive interaction.

However, the above argument will run into difficulty when we go to the strong interaction region, where we have to consider the effect of Anderson localization[3], so Anderson's "dirty superconductor" theory is only correct when the disorder in the system is not too strong.

Anderson Localization

The traditional view about the effect of disorder is that scattering by random potentials causes the Bloch waves to lose phase coherence over the length scale of the mean free path ℓ, but the wavefunction still can extend throughout the sample. It is Anderson who pointed out in his 1958 paper[3] that, the electronic wavefunction in a random potential may be profoundly altered if the randomness is sufficiently strong, for example, the wavefunctions may become localized instead of the traditional picture. In the localized region, the electronic wavefunction decays exponently, and looks like

$$| \psi(\mathbf{r}) | \sim \mathbf{exp}(| \mathbf{r} - \mathbf{r_0} | /\xi), \tag{1}$$

where ξ is called the localization length. This is because, at the limit of very strong disorder, even though orbits that are nearby in space still overlap significantly, they are in general quite different in energy due to the deep fluctuation in the random potential introduced by the disorder. According to the tight-binding theory, only near-degenerate states can bind together to form bands, and extend throughout the sample. As a consequence, the admixtures between different orbits can not propagate very far; on the other hand, states that are near degenerate are in general very far apart in space, so the overlap between them is extremely small.

Why 2D?

First of all, as it will be shown later, vortex and anti-vortex pairs play an important role in the superconductor-insulator transition. In 2D, the attractive interaction between a pair of vortices in a superconducting film falls off like $1/r^2$ at large distances[4][5]. As a result, vortex and anti-vortex pairs can happen even without the presence of external magnetic field. In addition, since the sheet resistance of a square film is proportional to its length and inversely proportional to its width, it should be independence of the size of the film, but not of the film's thickness. Hence, the thickness of the superconducting films is the only parameter we need to adjust to change the disorder in the films, and this is much easier to realize experimentally. In 3D, it is hardly possible to extract truely the influence of disorder on the disappearance of superconductivity, because by changing the composition one changes the electron liquid parameters as well. As a result, results from 2D films are more convincing and conclusive. Finally, in contrast to three-dimensional samples where a critical disorder is needed to localize the electronic states, 2D superconducting films at sufficiently low temperatures are localized no matter how weak the disorder is[6], hence in 2D, we have one parameter less to worry about.

2 Quantum Phase transitions

A quantum phase transition(QPT) is a transition taking place at the absolute zero of temperature. It is accomplished by changing not the temperature, but some parameters in the Hamiltonian of the system. As a result, the quantum ground state changes when the system crosses the phase boundary[6].

Quantum phase transitions are different from non-zero temperature transitions, or classical phase transitions(CPT), because quantum effects are central to QPTs, but not to CPTs, at least near critical points. According to the scaling hypothesis, the long-range correlation of fluctuations near $\mathbf{T_c}$ is responsible for all the physics of critical phenomena, in other words, the detail over smaller scales does not account for all the singular behavior of the system near $\mathbf{T_c}$. In CPTs, sufficiently close to the critical point, quantum fluctuations are important at the microscopic scale but not at the larger scales that control the critical behavior. It is classical thermal fluctuations that rule at long wavelengths. On the other hand, for QPTs, where $\mathbf{T_c}$ =0 , the critical fluctuations must be treated quantum mechanically.

d+1 Dimensional XY model(QPT at T=0)

Solving problems exactly is usually very difficult, and not necessary, if we care, for instance, only about whether there exists a superconductor-insulator transition(SIT) in 2D systems. To avoid solving exactly the ground states of a d-dimensional quantum system at the absolute zero of temperature, we can map it into a d+1 dimensional classical system with a non-zero fake temperature which is some measure of zero point fluctuations in the quantum system.

Inspired by Feynman's path-integral formulation of quantum mechanics that the net transition amplitude between two states of the system can be calculated by summing over amplitudes for all possible paths between them, we can express the partition function Z of the d-dimensional quantum system in the following way

$$Z(\beta) = \sum_n \langle n|e^{-\beta H}|n\rangle. \tag{2}$$

By taking $N\delta\tau = \hbar\beta$, we then have

$$Z(\beta) = \sum_n \sum_{m_1,m_2,\ldots,m_N} \langle n|e^{-(1/\hbar)\delta\tau H}|m_1\rangle\langle m_1|e^{-(1/\hbar)\delta\tau H}|m_2\rangle\langle m_2|\ldots|m_N\rangle\langle m_N|e^{-(1/\hbar)\delta\tau H}|n\rangle \tag{3}$$

It is clear now that calculating the thermodynamics of a quantum system is the same as calculating transition amplitudes for its evolution in imaginary time. In addition, if we think of the imaginary time as an additional spatial dimension, and also since the size of the extra dimension, $\hbar\beta$, diverges as T $\to$ 0, we get a truly d+1 dimensional effective classical system. The reason why we can call this d+1 dimensional system classical is that, we have already used up the physical temperature to define the extra spatial dimension, as a consequence, we have to find something else to act as a fake temperature K of this fictitious d+1 dimensional system, which is generally not zero.

So far we have established the equivalence between a quantum system and a classical system with an extra " temporal" dimension. If we can show that a phase transition exists in the fictitious classical system, it must have a mirror image in the corresponding quantum system. This is what makes d+1 dimensional XY model so beautiful.

In the vicinity of a continuous QPT, the correlation length diverges as the transition is approached. As a result, it should take a longer and longer time to propagate information across the distance of the correlation length. Therefore, in the d+1 dimensional classical system, besides the spatial correlation length ξ, we also need to define ξ_τ as the correlation "length" in the time direction. Generally, as T $\to$ 0, ξ and ξ_τ have the following singular behavior as δ goes to zero,

$$\xi \sim |\delta|^{-\nu}, \xi_\tau \sim \xi^{z}. \tag{4}$$

where δ is the difference between the tuning parameter(for instance, in the case of SIT, film thickness or magnetic field) and its critical value, and ν and z are called the correlation length exponent and the dynamic critical exponent, respectively.

Since, close to the critical point, there is no characteristic length scale other than ξ itself and no characteristic time scale other than ξ_τ, physical quantity O in the critical region near the transition should have the following scaling form

$$O(k, \omega, K) = \xi^{d_O} O(k\xi, \omega\xi_\tau), \tag{5}$$

where d_O is called the scaling dimension of the observable O measured at wave vector k and frequency ω; at the scale-invariant critical point, both ξ and ξ_τ diverge to infinity, in order for the observable to have a sensible limit as the critical point is approached, the scaling form should be correspondingly adjusted to the following form

$$O(k, \omega, K_c) = k^{-d_O} \tilde{O}(k^z/\omega). \tag{6}$$

The above mentioned diverging lengths and the associated scaling of physical properties are particularly important, because they represent the universal behavior which is insensible to microscopic details within certain global constraints as symmetry and dimensionality.

d+1 dimensional XY model can be very useful in proving the existence of QPTs by using computer simulations.

T $\neq$ 0: Finite Size scaling

The discussion above would be more powerful and practical if we knew how to do experiments at zero temperature to get data and plot out the scaling relations. Inevitably, we have to answer the following questions: Can we learn anything at all about QPTs by studying the T $\neq 0$ problem? If the answer is positive, what regions should we probe, and how should we modify the scaling forms?

Above some critical temperature, the system will become disordered due to huge thermal fluctuations; just in the same way, if we make K, the fake temperature large, the quantum fluctuations can then be so large that the system cannot order at any temperature. As a result, phase transitions can happen only if T and K are not very large. Specially, in order to observe QPTs, T must be close enough to the absolute zero so that some quantum effects will not be totally wiped out by thermal fluctuations.

T $\neq 0$ means that the size of the temporal dimension $L_\tau = \hbar\beta$ is finite. There is no definite answer on what the dimensionality of our system has, depending on how large a length scale we are interested in. Therefore, when we tune the physical temperature from zero to infinity, the dimensionality of our system can be changed continuously from d+1 to d.

Unlike in QPTs where we move the system away from the critical fixed point by tuning a parameter in the Hamiltonian, we do so here by reducing its dimensionality.

Suppose we fix the fake temperature K at somewhere pretty close to K_c, and we therefore assume that the behavior of the system is controlled by the same physics as at the critical point K_c. Corresponding to this fixed fake temperature, the correlation length ξ_o in a d+1 dimensional classical system(the mirror image of a d dimensional quantum system at T=0) is determined. When we decrease L_τ until it is comparable to ξ_o, a quantum-to-classical crossover will happen there, because, no matter how much we further reduce L_τ, the system could not tell the difference(since the correlation lengths are greater than the size of the system now.). We can also try to understand this quantum-to-classical crossover from a different approach. Corresponding to $\xi_{\tau o}$, there is a temperature T_o. When $k_B T_o/\hbar$ exceeds all the quantum fluctuation frequencies of interest, thermal fluctuations will bury all the effects of "long-wavelength" quantum fluctuations, so the critical fluctuations will behave classically; If the characteristic fluctuation frequencies obey $\hbar\omega \gg k_B T_o$, quantum fluctuations will be dominant. As a consequence, there exists a quantum-to-classical crossover, and if we have predicted the phase diagram correctly, the existence of QPT and this crossover near the critical point should be of mutual dependence. But, before we make any judgement about QPT, we must make sure that the crossover we have observed does indeed happen in the critical region. The formal theory of the effect of reduced dimensionality near critical points is called finite-size scaling. It says, when $|K - K_c|/K_c \ll 1$ and T $\to 0$, physical qualities have the finite scaling form

$$O(k, \omega, K, T) = {L_\tau}^{d_O} O(k {L_\tau}^{1/z}, \omega L_\tau, L_\tau/\xi_\tau). \tag{7}$$

In the case of SIT, the correlation lengths still have the same diverging properties as

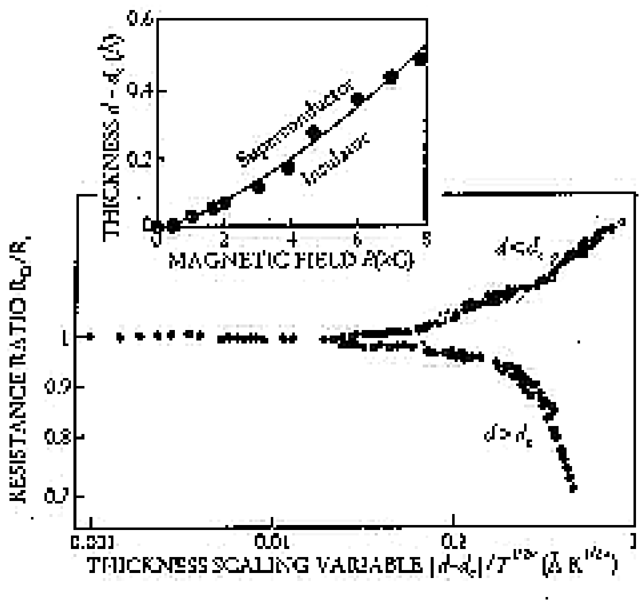

Figure 1:

in (4). The resistance of a 2D superconducting film follows the scaling relation

$$R(\delta, T) = R_c f(\delta T^{-1/\nu z}), \tag{8}$$

where R_c is the critical resistance at δ=0, and f(x) is a universal function such that f(0)=1. Since, in the critical region, the correlation lengths go to infinity, we get the same physics even by changing the spatial resolution of the system, as long as this spatial resolution is much less than the correlation lengths. Hence, we should observe a collapse of data whenever a transition happens, which will be the single tool we will use to make any judgement about whether a transition occurs.

The superconductor-insulator transition can be realized by tuning either the thickness of the films or the magnitude of external magnetic fields. FIG 1[8], 2[9] and 3[10] clearly indicate the existence of a phase transition. If the transition indeed happens in the quantum critical region, we can then conclude that the superconductor-insulator transition at T=0 is a quantum phase transition.

3 Different Models

Various models of the superconductor-insulator transition in disordered films can be divided into two groups: those in which the superconductivity is destroyed by solely phase fluctuations of the order parameter, and those which focus on the amplitude

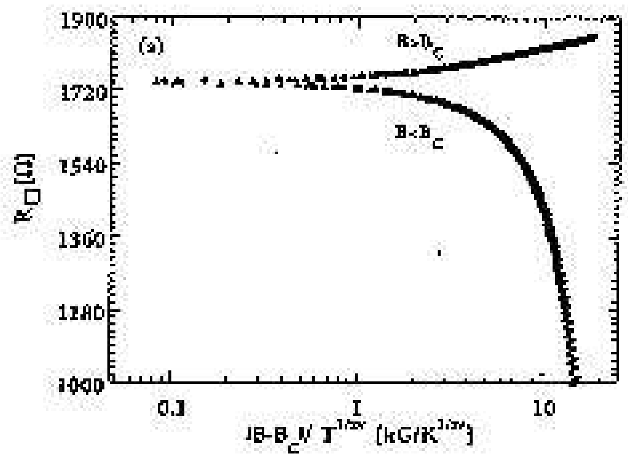

Figure 2:

fluctuations of the order parameter. However, the different interpretations agree on one point that, Cooper pairs, vortex and anti-vortex pairs, and disorder play the most important roles in the SIT.

Kosterlitz-Thouless-Berezinski Transition[11][5]

In 1935, Peierls has argued that thermal motion of long-wavelength phonons will destroy the long-range order of a two-dimensional solid in the sense that the mean square deviation of an atom from its equilibrium position increases logarithmically with the size of the system. However, as pointed out by Kosterlitz and Thouless[11], in two-dimensional system, there may exist an nonconventional long-range order, namely topological long-range order.

Although free vortices cannot occur at low temperature in a large system since their energy increases logarithmically with the size of the system, as has been shown by Pearl[4], the force of attraction between a pair of vortex-antivortex in the superconducting film is proportional to $1/r^2$ at large separations. Consequently, vortex-antivortex pairs must occur because of thermal fluctuations. However, on the other hand, when the temperature is high enough, the largest pairs are unstable. As we increase the temperature to Tc, topological long-range order up to a length scale L_s, which is the effective London penetration length, will be destroyed. It is believed that there is a phase transition at that point, which is called Kosterlitz-Thouless-Berezinski transition(KTB).

One should expect to see a KTB transition except there are sample-dependent length scale due to granularity or defects that cut off the long-range vortex interaction. So the films we are interested to explore should, first of all, be homogeneous, which is

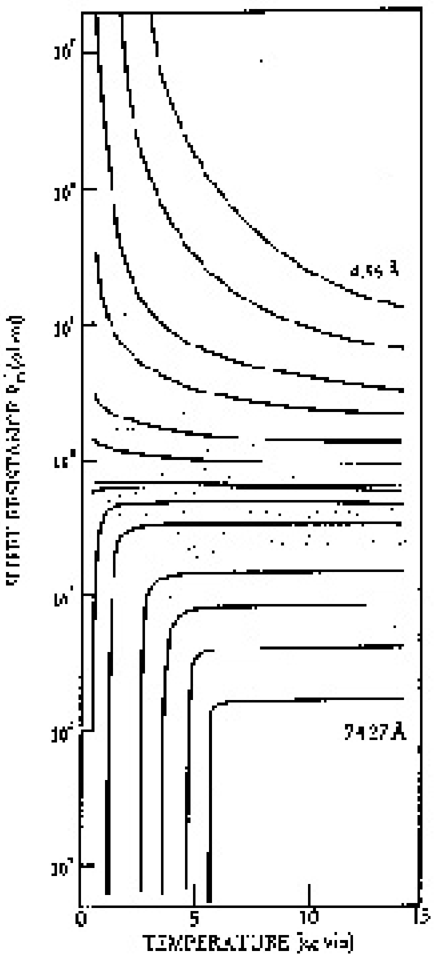

Figure 3:

required also because, if the films are granular, the Josephson coupling between grains of such films will be determined by a critical value of the electrostatic energy to the Josephson coupling energy, which in general depends on the geometry and is generally hard to control; secondly, the films should not be too thin or disorder so that electrons are strongly localized, but thinner than the effective London penetration length so that we really get a 2D film in the sense that KTB is allowed to happen.

Duality Transformation and Dirty Boson Model

Duality transformations[12] are mathematical methods that relate the physical properties of one system to those of another physical system. If such a transformation exists, the systems are said to be dual to each other. For example, in electromagnetism, a duality transformation interchanges the role of the electric and the magnetic fields, and electric charges and magnetic monopoles. In another example[13], a remarkable symmetry has been observed between the diagonal, nonlinear, current-voltage($I - V_{xx}$) characteristic taken in the fractional quantum Hall effect (FQHE) liquid state of the two-dimensional electron system and those taken in the bordering insulating phase. With the role of currents and voltages exchanged, the systems before and after the transformation look identical, which can be interpreted as evidence of the the existence of charge-flux duality symmetry in the system. We are interested in 2D superconductors with both Cooper pairs and Vortices coexisted. Since vortex in 2D is point defect, it is possible to perform dual transformation that will interchange the roles of Cooper pairs and vortices. Vortices are topological point defects in the condensate wavefunction, hence every time a Cooper pair moves around a vortex, the quantum phase of the system should change by $\pm 2\pi$. Associated with the gradient in the phase is a circulating current, much like a whirlpool in an ordinary fluid. In the spirit of duality, we can also view vortices as the fundamental degrees of freedom instead of Cooper pairs, and we can consider the wavefunction for the system depends actually on the positions of the vortices. Therefore, each time, a vortex moves around a Cooper pair, the phase of the vortex wavefunction winds by $\pm 2\pi$. As shown in FIG 4, Cooper pairs and vortices are moving perpendicularly to one another. The electric current along the sample is given by $I = (2e\dot{N}_c)$, where $\dot{N}_c$ is the flux of Cooper pairs; the voltage drop between the two ends of the sample is given by $V = (h/2e)\dot{N}_v$, where $\dot{N}_v$ is the flux of vortices moving transverse to the current.

Superconductor-insulator transition can be driven by either tuning the thickness of the 2D film or applying and varying a magnetic field which is perpendicular to the film surface. In a pure system the vortices will freeze into an Abrikosov vortex lattice below a melting line B_m as shown in FIG 5[14]. In a real system with disorder, the long-range correlations will be destroyed. At finite temperature, vortex creep will destroy phase coherence and lead to a resistance. At T=0, however, due to

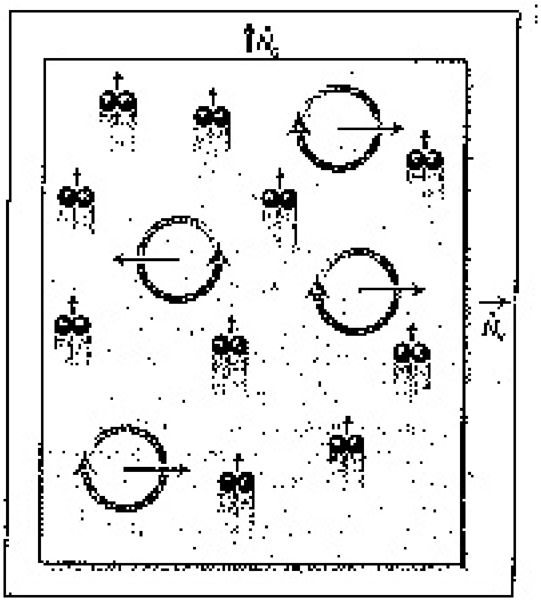

Figure 4:

disorder, the vortex and anti-vortex pairs can be localized, even though the existence of quantum fluctuations. Therefore, the vortex flux and hence the voltage will vanish, meaning that the current flows without dissipation. This is called a vortex glass, and it is a superconducting phase. As we increase the magnetic field, there exists another possibility. The density of the vortex and anti-vortex pairs will increase with the increase of the field, and the pairs should delocalize and undergo a Bose condensation at some critical field. Contrary to the vortex glass phase, here the Cooper pairs are localized, and we get a zero current flow. This is called Bose insulator phase, and it is the dual of the vortex glass phase. At the critical point where the system is switching from the vortex glass to the Bose insulator, both charges and vortices are simultaneously mobile, and hence the system should have finite dissipation. As a result, the resistance at the critical point should be $R^* = h/(2e)^2(\dot{N}_v/\dot{N}_c)$. Suppose the 2D superconductor system is self-dual, in other words, the dual system is identical to the original system, then the vortices and charges would necessarily behave in the same way at the critical point. We should then expect $\dot{N}_v = \dot{N}_c$, and therefore the resistance at the critical point should be universal, namely $R^* = h/(2e)^2$. It is clear that, the Dirty Boson model has assumed that superconductivity is destroyed only by phase fluctuations of the order parameter, and Cooper pairs persist on the insulating side of the transition. One importance prediction of the Bosonic model is the universal value of theoretical resistance. But, this is not the whole story yet. Paalanen, Hebard and Ruel[15] have studied the magnetoresistance and the Hall effect in amorphous

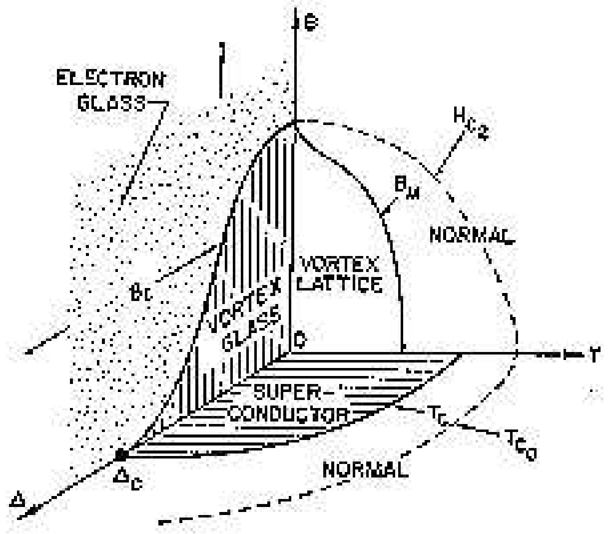

Figure 5:

InO_x films and observed two distinct transitions: one at the a critical field ${B_{xx}}^c$ where the longitudinal resistance diverges and the system presumably undergoes a transition from a superconducting vortex glass to a Bose insulator with localized Cooper pairs as shown in FIG 6, and the other at a higher field ${B_{xy}}^c$, where the transverse resistance diverges, and it is believed that the Cooper pairs in the Bose insulator begin to unbind. The new insulating phase is called the Fermi insulator. The transition, as shown in FIG 7[15], in the transverse resistance occurred at the same magnetic field where the longitudinal resistance showed a maximum. Since a Bose insulator might be expected to have a higher resistance than a Fermi insulator, it is therefore evident that the magnitude of the order parameter persists across the superconductor-insulator transition.

A Fermionic Picture of SIT

Although,the Dirty Boson model has successfully explained many experimental results, we must realize that the Bosonic picture is not completely correct, at least, it does not take into account the role of unpaired electrons which would be expected to become important when either the magnetic field or the disorder are high enough to ensure the dissolution of the localized Cooper pairs. What's more, experiments have shown a lack of a universal critical resistance, and it was also found out that the superconducting gap and the mean field transition temperature are both suppressed as disorder is increased, and that the gap vanishes on the insulating side of

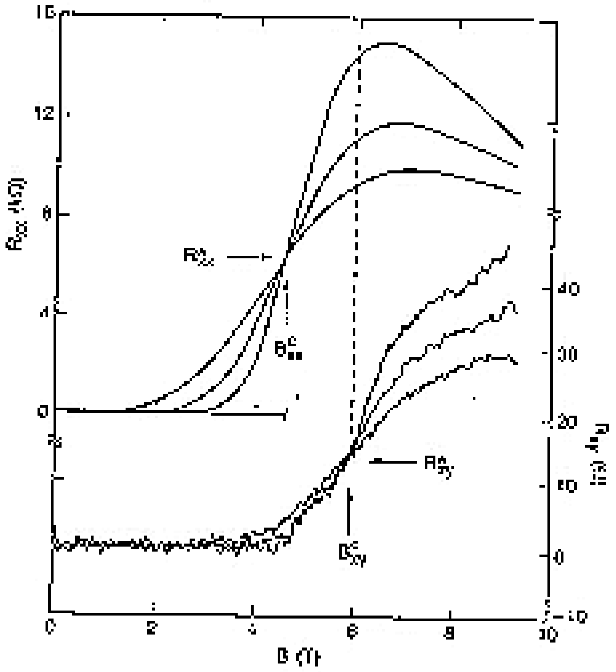

Figure 6:

the superconductor-insulator transition[16].

An alternative picture of interacting electrons propose a different mechanism[17][18][19]: the density of states and the Cooper pairs are suppressed on the insulating side of the Superconductor-insulator transition due to an enhenced Coulomb interaction. The SIT occurs as a consequence of fluctuations in the amplitude, rather than the phase of the order parameter. In other words, Cooper pairs break up into single electrons at the transition. Therefore, the superconducting gap should also vanish at the transition.

This Fermionic picture seems to be able to compensate some shortcomings of the Bosonic description. This may imply that a full picture must take all of them into account.

4 References

[1] A. Goldman, N. Marković, Physics Today, November, 391998.
[2] P.W. Anderson, J. Phys. Chem. Solid 11, 26(1959).
[3] P.W. Anderson, Phys. Rev. 109, 1492(1958).
[4] J. pearl, Appl. Phys. Lett. 5, 65(1964).

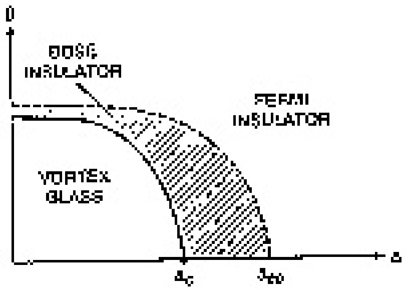

Figure 7:

[5] B.I. Halperin and David R. Nelson, J. Low Temp. Phys. 36, 599(1979).
[6] E. Abrahams, P.W. Anderson, D.C. Licciardello, and T.V. Ramakrishnan, Phys. Rev. Lett. 42, 673(1979).
[7] S. Sondhi, S. Girvin, J. Carini, D. Shahar, Rev. Mod. Phys. 69, 315(1997).
[8] N. Marković, C. Christiansen, G. Martinez-Arizala, A. Mack, A. Goldman, Phys. Rev. Lett. 74, 2331(1995).
[9] A. Yazdani, A. Kapitulnik, Phys. Rev. Lett. 74, 3037(1995).
[10] Y. liu, K. McGreer, B. Nease, D. Haviland, G. Martinez, J. Halley, A. Goldman, Phys. Rev. Lett. 67, 20681991.
[11] J.M. Kosterlitz, D.J. Thouless, J. Phys. C 6, 1181(1973).
[12] S.M. Girvin, Science 274, 524(1996).
[13] D. shahar, D.C. Tsui, M. Shayegan, E. Shimshoni, S.L. Sondhi, Science 274, 589(1996).
[14] M.P.A. Fisher, Phys. Rev. Lett. 65, 923(1990).
[15] M.A. Paalanen, A.F. Hebard, R.R. Ruel, Phys. Rev. Lett. 69, 1604(1992).
[16] J.M. Valles Jr., R.C. Dynes, L.P. Garno, Phys. Rev. Lett. 69, 3567(1992).
[17] S. Maekawa, H. Fuuyama, J. Phys. Soc. Japan 51, 1380(1981).
[18] D. Belitz, Phys. Rev. B 40, 111(1989).
[19] A. Finkel'shtein, JETP Let. 45, 46(1987).