Effect of Long Range Coulomb Interactions on the Mott Transition

R. Chitra and G. Kotliar

Serin Physics Laboratory, Rutgers University, Piscataway, New Jersey 08855

(Received 18 June 1999)

We reconsider the Mott transition problem in the presence of long range Coulomb interactions. Using an extended dynamical mean field theory (DMFT) that sums an important class of diagrams absent in ordinary DMFT, we show that in the presence of Coulomb interactions, the zero temperature Mott transition is, as envisioned by Mott, discontinuous in two and three dimensions.

PACS numbers: 71.30.+h

The correlation driven metal-insulator transition (MIT), i.e., Mott transition, is one of the most challenging problems in solid state physics. The nature of these transitions depends on the interplay of band structure, magnetism, and electronic correlations. The simplest scenario for a MIT was first put forth by Mott [2]. He suggested that a crystalline array of hydrogenic atoms, i.e., atoms with just one electron in their outermost shell, shows a zero temperature transition from a metal to an insulator as the density is decreased or, effectively, as the distance between atoms increases [3]. Later, other possibilities for the MIT, namely, Slater's band transition [4] where the system becomes insulating because of the doubling of the unit cell due to antiferromagnetism and the Brinkman-Rice scenario within the Hubbard model [5] which is characterized by a strong mass enhancement, were put forth.

Mott's original argument depended on the presence of Coulomb long range interactions and pointed out that since the number of "free" electrons could vary only discontinuously across the transition, the transition is necessarily first order. We recapitulate Mott's ideas below: the main premise was that, due to the long range nature of the interactions, the electrons and the holes in the two Hubbard bands will always form bound states and not exist independently. The condition for the formation of at least one bound state, however, depends on the screening of the Coulomb interaction. Using Thomas-Fermi estimates for screening by N electrons per unit volume, Mott showed that no bound states exist provided

$$N^{1/3} a_H > 0.4\,, \tag{1}$$

where a_H is the Bohr radius. In other words, the system is metallic as long as the density satisfies (1). For densities where (1) is violated, the electrons and holes in the Hubbard bands always form bound pairs resulting in an insulating behavior, implying that the number of carriers jumps at the transition. Since this results in a kink in the free energy, the transition is discontinuous and first order. For an illuminating discussion of these ideas see Refs. [6,7].

On the other hand, various studies of the Hubbard model [3,5,8] which has long been used as the prototype to describe various aspects of real systems [1] indicate that the zero temperature transition is continuous. This model contains a term which describes the hopping of electrons between different atomic sites and another describing the Coulomb repulsion felt by two electrons on the same atom. In the past few years, the nature of the MIT in this model has been clearly elucidated using the dynamical mean field theory (DMFT) [8]. It was shown that the zero temperature MIT was characterized by a continuous vanishing of a Kondo-like resonance at the Fermi level in the metal at the transition, leading to an insulator with a preformed gap. This led to a divergence of the effective mass of the quasiparticles and hence the linear coefficient of the specific heat as the transition is approached from the metallic side [5,8].

The purpose of this work is to incorporate some of the effects of the long range Coulomb interactions and Mott's ideas into the framework of dynamical mean field theory. For this purpose, we explore the effects of the long range interaction on the MIT seen in the single band Hubbard model [3], using a simple extension of DMFT [11,12]. It was shown, in a model of spinless fermions, that this approach captures important $\frac{1}{d}$ corrections [11]. This method was independently developed in Ref. [12] and it was applied to the problem of the breakdown of Fermi liquid theory. We first describe the approach by isolating a class of diagrams which can be formally controlled by scaling the interactions and the kinetic energy appropriately and which can be summed using impurity models. We then demonstrate that when this extended DMFT is applied directly to a three- or two-dimensional lattice with interactions having the Coulomb form, it changes the Mott transition which was continuous in ordinary DMFT, to a discontinuous first order transition, reiterating Mott's ideas.

The effective Hamiltonian we use to describe our system is a generalization of the Hubbard Hamiltonian:

$$H = \sum_{\langle ij\rangle\sigma} t_{ij} c^{\dagger}_{i\sigma} c_{j\sigma} + \text{H.c.} + \sum_{i\sigma} \mu c^{\dagger}_{i\sigma} c_{i\sigma} + \sum_{i} U n_{i\uparrow} n_{i\downarrow} + \sum_{i\neq j,\sigma\sigma'} V_{ij} : n_{i\sigma} n_{j\sigma'} :. \tag{2}$$

The first term is the hopping matrix element for an electron from site i (representing the ion at R_i) to its neighboring site j, μ is the chemical potential, U is the Coulomb repulsion felt by the electrons when they are on the same atom, and the normal ordered last term is the Coulomb

 0031-9007/00/84(16)/3678(4)$15.00

interaction between electrons on different atoms. The coupling constants are given by overlap integrals involving a set of chosen basis vectors $u(r - R_i)$ for, e.g., the Wannier or Hartree basis. For example, the on-site interaction U is given by

$$U = \int dr\, dr' |u(r)|^2 |u(r')|^2 \frac{e^2}{|r - r'|}$$
$$V_{ij} = \int dr\, dr' |u(r - R_i)|^2 |u(r' - R_j)|^2 \frac{e^2}{|r - r'|}. \quad (3)$$

The Hamiltonian (2) is now studied in the dynamical mean field approximation. The first step in the DMFT is to scale the parameters in the large d limit appropriately such that the corresponding energy terms remain finite. In addition, the scaling should be chosen such that the terms of interest remain relevant in the large d limit. In the conventional DMFT scheme, the interaction is not scaled. Consequently, all nonlocal interactions can be neglected in the large d limit. Here, to manifestly retain the effects of the nonlocal terms, we adopt the following scaling: t_{ij} is scaled as $\sqrt{d}^{-|i-j|}$ ($t \to \frac{t}{\sqrt{d}}$ in the case of nearest neighbor sites); $U \to U$ and V_{ij} are scaled by $\sqrt{d}^{-|i-j|}$. The normal ordering in (2) makes this scaling well defined and the leading diagrams can be summed up using the cavity method described in Ref. [8]. It is a prescription for choosing a set of diagrams that contribute to the self-energy and hence the Green's function. DMFT prescribes a scheme for choosing different sets of diagrams that contribute to the different Green's functions. In the case of the Hubbard model ($V = 0$), this corresponds to retaining only skeleton diagrams constructed from U and the local Green's functions G_{ii} in the self-energy, resulting in the self-energy Σ_{ij} being local. In the presence of the longer range interaction, the above scaling retains all skeleton diagrams constructed using the local Green's function G_{ii} and the interaction vertices U and V_{ij}, such that every point i which has a vertex V_{ij} originating from it has another vertex V_{ki} terminating at it. In effect, this corresponds to replacing the U by an U_{eff} in the local self-energy evaluated for a Hubbard model. An example of the diagrams retained is shown in Fig. 1. We mention that the Hartree term (arising only from U) is generated in the cavity method. Depending on the basis chosen to derive the effective model parameters (2) care should be exerted to see that the Hartree term is not double counted. We neglect the Fock term in the ensuing calculations, since the Fock term is of higher order in $\frac{1}{d}$ than the Hartree term.

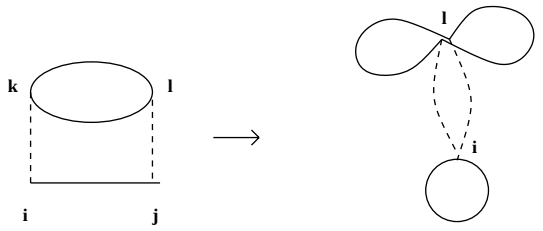

FIG. 1. Reduction of the self-energy diagrams in the limit of infinite d.

Using the cavity method [8] and integrating out all sites save a chosen site or cavity o, we obtain the following local effective impurity action with retarded interactions:

$$S_{\text{eff}} = \int d\tau\, d\tau' \sum_{\sigma} c^{\dagger}_{0\sigma}(\tau) \mathcal{G}_0^{-1}(\tau - \tau') c_{0\sigma}(\tau') + U n_{0\uparrow}(\tau) n_{0\downarrow}(\tau) - \sum_{\sigma\sigma'} n_{0\sigma}(\tau) \Pi_0^{-1}(\tau - \tau') n_{0\sigma'}(\tau'), \quad (4)$$

where

$$\mathcal{G}_0^{-1} = \partial_\tau - \sum_{ij} t_{0i} t_{0j} G^{(0)}_{ij}(\tau - \tau') \quad (5)$$

and the retarded interaction

$$\Pi_0^{-1}(\tau - \tau') = \sum_{ij} V_{0i} V_{0j} \Pi^{(0)}_{ij}(\tau - \tau') = \sum_{q} V(q)^2 \Pi^{(0)}(q, \tau - \tau') \quad (6)$$

with $G^{(0)}_{ij}$ denoting the single particle Green's function and $\Pi^{(0)}_{ij} = \langle \sum_{\sigma\sigma'} n_{i\sigma}(\tau) n_{j\sigma'}(\tau') \rangle_0$. The superscript (0) implies that the quantities have been evaluated in the system from which the site 0 and all the links to this site have been removed. These in turn are related to quantities evaluated on the entire lattice. Summing over the sites leads to $\mathcal{G}_0$ and Π_0 being determined by self-consistent equations involving local quantities evaluated on the full lattice. Note that the effect of all nonlocal quartic interactions is to dynamically screen the on-site repulsion, with the screening potential given by (6). These equations depend on the nature of the lattice used. For example, on the Bethe lattice, $G^{(0)}_{ij} = G_{ij}$ and $\Pi^{(0)}_{ij} = \Pi_{ij}$. Retaining only the on-site and nearest neighbor interactions, the self-consistency condition that should be satisfied by the retarded interaction on the Bethe lattice takes the simple form $\Pi_0^{-1} = V^2 \Pi_{\text{loc}}$, where Π_{loc} is the local density density correlator. For arbitrary lattices and interactions (i.e., general t_{ij} and V_{ij}) the (extended) dynamical mean field equations can be derived by generalizing the discussion presented here to the case where the hoppings t_{ij} are scaled by $\sqrt{d}^{-|i-j|}$.

To obtain the Green's functions of (4), we now define certain irreducible quantities. The lattice Green's functions can be expressed in terms of a self-energy Σ which is one particle irreducible and which becomes local in the limit of infinite dimensions.

$$G(i\omega_n, q) = \frac{1}{i\omega_n - \epsilon_q - \Sigma(i\omega_n)}, \quad (7)$$

where ϵ_q is the dispersion on the lattice and $i\omega_n$ are the Matsubara frequencies. Similarly, the density density correlator $\Pi(q, i\omega_n)$ on the lattice defines an irreducible part $\tilde{\Pi}$ via the Dyson equation

$$\Pi(q, i\omega_n) = \frac{\tilde{\Pi}}{1 + V(q)\tilde{\Pi}}, \tag{8}$$

where $\tilde{\Pi}$ is the sum over all polarization diagrams constructed with the full G and interaction lines $V(q)$ such that all the diagrams are irreducible with respect to $V(q)$. Within the infinite d approximation, since all vertex functions become independent of momenta, $\tilde{\Pi}$ also becomes independent of momenta. The impurity model (4) allows us to compute all local quantities, in particular, Σ and $\tilde{\Pi}$ and hence, the local density density correlator $\Pi_{\rm loc}$ and the local one particle Green's function $G_{\rm loc}$ as functionals of the Weiss fields $\mathcal{G}_0$ and Π_0.

$$\Pi_{\rm loc}^{-1}\{\Pi_0, \mathcal{G}_0\}(i\omega_n) = \tilde{\Pi}^{-1}\{\Pi_0, \mathcal{G}_0\}(i\omega_n) - \Pi_0^{-1}(i\omega_n) \tag{9}$$

and

$$\Sigma\{\Pi_0, \mathcal{G}_0\}(i\omega_n) = \mathcal{G}_0^{-1}(i\omega_n) - G_{\rm loc}\{\Pi_0, \mathcal{G}_0\}(i\omega_n). \tag{10}$$

Using (8) and (9), we can eliminate $\tilde{\Pi}$ to obtain a self-consistent equation for $\Pi_{\rm loc} = \sum_q \Pi(q, i\omega_n)$,

$$\Pi_{\rm loc}\{\Pi_0, \mathcal{G}_0\} = \sum_q \frac{1}{\Pi_{\rm loc}^{-1}\{\Pi_0, \mathcal{G}_0\} - \Pi_0^{-1} + V(q)}. \tag{11}$$

Similarly, we obtain an equation for $G_{\rm loc}$

$$G_{\rm loc}\{\Pi_0, \mathcal{G}_0\} = \sum_q \frac{1}{i\omega_n - \epsilon_q - \Sigma\{\Pi_0, \mathcal{G}_0\}}. \tag{12}$$

These equations were derived by an appropriate scaling of the interactions and the hopping elements in the original model so as to obtain a well-defined limit of large coordination. In the spirit of DMFT [8], we can, however, regard these equations as defined on a finite dimensional lattice by replacing V by the usual Coulomb interaction on the lattice. We use these equations to make qualitative predictions of the effect of the Coulomb long range interaction in finite dimensions, on the order of the Mott transition seen.

The continuous character of the Mott transition at zero temperature signaled by the vanishing of a Kondo-like resonance at $\omega = 0$ at U_{c2} was established using the projective self-consistent method within DMFT [9]. We briefly review the projective method below. Then, using the results of Ref. [9] we study how long range Coulomb interactions treated within the extended DMFT modify the above transition.

The projective method uses the separation of two energy scales that exists in the metallic phase close to the MIT, i.e., the low energy scale wD (where w and D are the weight of the Kondo resonance seen in the metal and the half bandwidth, respectively) and the high energy scale of the Hubbard bands, U. w is related to the quasiparticle residue $z = w/1.7$ [9]. The high energy scales are then eliminated to obtain an effective theory governed by the low energy scale w, which goes to zero at the MIT [8]. Using this effective low energy model, which is a Kondo model of an impurity spin interacting with a bath of electrons, we can obtain the free energy or the ground state energy of the lattice problem close to the transition point. This free energy correctly describes the low energy and coherent part of the spectra. To order w, the low energy Kondo problem has been derived in Ref. [9] and is given by

$$\mathcal{H} = w\Gamma \mathbf{S} \cdot \mathbf{s}_L + \mathcal{H}_b. \tag{13}$$

Here $\mathcal{H}_b$ describes a band of low energy conduction electrons and $\mathbf{s}_L$ represents the local spin operator of these electrons. $\mathbf{S}$ is the impurity spin and the Kondo coupling Γ is determined by matrix elements in the high energy sector (comprising the two Hubbard bands). The high energy sector is an insulator and has a spin doublet ground state which is separated from the excited states by a large gap. Using the results of Ref. [9], the simplest approximation to the high energy sector of the Hubbard model yields $\Gamma = \frac{2D^2}{U}$. The self-consistency conditions translate to conditions on the expectation value $\langle \mathbf{S} \cdot \mathbf{s}_L \rangle$. Since we are interested only in the qualitative features of the transition, we replace the bath of electrons by a single electron. Therefore, using this toy model, the energy of the Kondo model is

$$E_K = -\frac{3}{4} w\Gamma. \tag{14}$$

Using the results of [10], and taking into account the kinetic energy of the lattice which is positive and of order wD, the ground state energy of the lattice model has the following expansion in terms of w:

$$E_0 = \frac{D}{2}\left(1 - \frac{U_c}{U}\right)w + \frac{3D^2}{8U} w^2 + \text{corrections}. \tag{15}$$

The w^2 term describes the lowest order effective interaction between the bath electrons. The transition is signaled by the vanishing of the term linear in w in (15) at $U = U_c = 3D$. Since the energy minimum still occurs at $w = 0$, this transition is continuous.

Though the DMFT and the self-consistent projective method are both formulated in infinite dimensions, it is nonetheless known that they capture some aspects of the physics of systems in finite dimensions rather well. We therefore use the projective method to make some predictions about the transition in three dimensions. First, in the presence of the long range Coulomb interaction, there are corrections to (15), because of the screening of the on-site Coulomb interaction given by (6). This implies that U in (15) has to be replaced by an $U_{\rm eff}$. For the free energy (15), it is sufficient to consider the effective static on-site repulsion. Assuming $\Pi^{(0)} = \Pi$ + sub-(leading corrections) in (6) and using the fact that at zero frequency, $\tilde{\Pi}$ in (8) is by definition the compressibility κ of the system, one obtains from (6) and (4)

$$U_{\rm eff} = U - \sum_q V^2(q) \frac{\kappa}{1 + \kappa V(q)}. \tag{16}$$

Note that $V(q) = C/(qa)^2$ is the Fourier transform of the Coulomb potential with strength C. Since the effective interaction remains finite at the transition, the leading critical behavior of the compressibility κ is the same as that found in the Hubbard model where [8,9]

$$\kappa \simeq 0.55 \frac{w}{D} . \tag{17}$$

Within the Hubbard model, κ is finite in the metal and goes to zero smoothly as $w \to 0$ at the transition. Notice that the parameter w in the projective self-consistent method neatly embodies the notion of "a number of free carriers" in Mott's original work. Equation (16) shows that the effective repulsion seen by the electrons in the metallic phase is reduced. In the insulator $\kappa = 0$ and the bare on-site repulsion is unscreened. This is physically correct because one does expect the itinerant electrons in the metallic phase to better screen the Coulomb repulsion. Using (17) in (16), and performing the q sum in $d = 3$ we find that the effective repulsion felt on-site is

$$U_{\text{eff}} = U - aw^{1/2}, \tag{18}$$

where $a = 0.157C\sqrt{C/D}$. The screening is proportional to $\log w$ in two dimensions. Replacing U by U_{eff} in (15), we see that the expansion of the energy in terms of the low energy weight w has nonanalytic terms

$$E_0 = \frac{D}{2}\left(1 - \frac{U_c}{U}\right)w - a\frac{3D^2}{2U^2}w^{3/2} + \frac{3D^2}{8U}w^2 + \text{correction}. \tag{19}$$

The nonanalyticity in (19) arises purely from the Coulomb long range interaction via the self-consistency conditions (11). We now see that at the transition, this nonanalyticity shifts the minimum of the energy which formerly was at $w = 0$ to a nonzero value. In other words, the value of w jumps discontinuously to zero at the transition, i.e., $\Delta w_{\text{trans}} = 9(\frac{a}{U_c})^2 = 0.216(C/U_c)^2(C/D)$ which equals $0.024(C/D)^3$ at $U_c \simeq 3D$. This implies a *minimal metallic quasiparticle residue* $z_{\text{min}} = 0.127(C/U_c)^2(C/D)$ and is a feature of a first order transition. In two dimensions, a similar analysis results in a $w \log w$ in the free energy (15) which also implies a discontinuous transition.

z_{min} depends on the magnitudes of the two independent model parameters, C and U which in turn depend on the material considered, and could in principle be determined from constrained density functional calculations. Normally, since the long range parts of the Coulomb interactions are much better screened than the on-site ones, we expect C to be smaller than U. As an order of magnitude estimate we take $C \approx D \approx U_c/3$ implies that $z_{\text{min}} \lesssim 0.1$. The jump in z_{min} will be more visible in systems with very poor screening.

To conclude, we have studied the effects of long range Coulomb interactions on the Mott transition using an extended DMFT. We find that the Coulomb interactions dynamically screen the effective on-site interaction. We find that to lowest order, the screening is determined by the low energy scale w which is related to the width of the Kondo resonance at the Fermi level. Consequently, the screening is zero in the insulator. Extending our analysis to three dimensions, we find that the screening term is nonanalytic in w resulting in nonanalytic terms in the free energy. As a result the Mott transition which was continuous in the absence of these dynamical screening terms now becomes a first order transition vindicating Mott's ideas. More importantly, the formalism developed in this Letter can be used to study a whole range where a more realistic treatment of the Coulomb interactions is required. Combined with a more realistic modeling of band structure and orbital structure, one could obtain estimates for z_{min} for real materials which undergo a pressure driven Mott transition.

[1] M. Imada, A. Fujimori, and T. Tokura, Rev. Mod. Phys. **70**, 1039 (1998).
[2] N. F. Mott, Philos. Mag. **6**, 287 (1961).
[3] J. Hubbard, Proc. R. Soc. London A **281**, 401 (1964).
[4] J. C. Slater, Phys. Rev. **82**, 538 (1951).
[5] W. F. Brinkman and T. M. Rice, Phys. Rev. B **2**, 4302 (1970).
[6] N. F. Mott, *Metal Insulator Transitions* (Taylor & Francis, London, 1974).
[7] C. M. Varma and T. Giamarchi, in *Strongly Interacting Fermions and High Tc Superconductivity,* edited by B. Doucot and J. Zinn Justin (Elsevier, New York, 1995).
[8] A. Georges, G. Kotliar, W. Krauth, and M. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996).
[9] G. Moeller, Q. Si, G. Kotliar, M. J. Rozenberg, and D. S. Fisher, Phys. Rev. Lett. **74**, 2082 (1995); G. Moeller, Ph.D thesis, Rutgers University, 1996.
[10] G. Kotliar (to be published).
[11] H. Kajueter and G. Kotliar (to be published).
[12] Q. Si and J. L Smith, Phys. Rev. Lett. **77**, 3391 (1996); J. Lleweilun Smith and Q. Si, cond-mat/9903083.