WAVES and SYMMETRY
Raman Centenary Symposium,
December 1988, Bangalore.

The Mott Transition

T. V. RAMAKRISHNAN
Department of Physics,
Indian Institute of Science, Bangalore 560 012, India.

I. INTRODUCTION

A. *Mott Insulator and Band Insulator*

The Mott transition[1] (or state localization by electron correlation) is a little more than half a century old as a distinctly recognized phenomenon; it is still not understood in detail. The phenomenon is now particularly relevant because of the recently discovered oxide superconductors. Strong correlation effects are quite important in these systems, perhaps central to their electronic behaviour. I present here a brief review and some ideas.

At a conference in 1937 organized by Mott, de Boer and Verwey[2] reported that NiO is a transparent insulator. This is rather unexpected, since according to the band theory of solids (developed in the late twenties) it ought to have been a half- filled band metal. The level scheme for the Ni d electron is shown in figure 1. In the cubic crystal field of NiO, the d level splits into t_{2g} (three-fold degenerate orbitally) and e_g (two-fold degenerate). The eight d electrons of Ni^{++} fill the t_{2g} band completely (six states per Ni^{++}) and the e_g band (four states per Ni^{++}) partially. The oxygen level derived band is completely occupied. Thus the substance ought to be, from this simple electron count, a metal. Instead, it is a good insulator with an energy gap of about 3 eV.

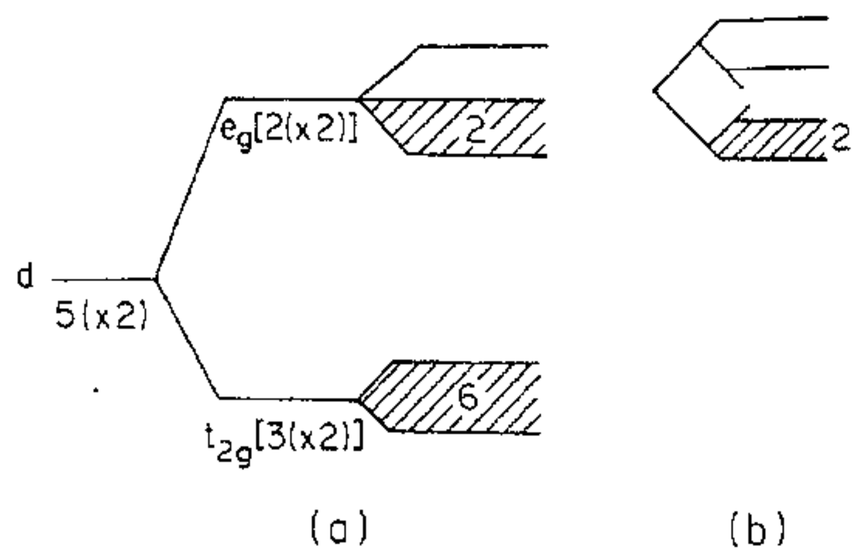

Figure 1. **(a)** Schematic illustration of how the degeneracy of d levels is lifted in an octahedrally-symmetric crystalline environment, and the resulting half-filled band (relevant for NiO). **(b)** Further lifting of degeneracy by two-sublattice magnetic ordering and formation of a semiconducting gap.

At the above conference itself, Peierls suggested that the insulating nature of NiO could be due to electron repulsion plus the Pauli exclusion principle discouraging configurations such as Ni^{+} and Ni^{+++}. The argument is simply that electron transport from one Ni^{++} site to another means that the former becomes Ni^{+} and the latter becomes Ni^{+++}. If one of these or both have a much higher energy than Ni^{++}, (because of extra coulomb energy and exclusion principle effects), electron

transport will require activation energy and the substance is an insulator. In the band model where interaction between electrons is ignored, different ionic configurations have, in effect, the same energy. Thus the NiO kind of insulator is solely due to correlation effects, and not due to crystal periodicity, band gaps and electron count per unit cell. The idea and its experimental consequences were developed by Mott[1] in a series of papers, reviews and books starting from 1937 (see also Refs. 3 and 4 for recent reviews). We will briefly mention some of these in this paper.

An alternate explanation for the insulating behaviour of NiO was provided by Slater[5] in 1951, after it was discovered that NiO is an antiferromagnet. The latter fact means that the (magnetic) unit cell is double the size of the chemical formula unit cell. The new unit cell has enough electrons to fully occupy the e_g-like one-electron bands; the new periodic potential due to antiferromagnetism splits the e_g band (see Figure 1). If this is indeed the reason why NiO is an insulator, one expects it to be metallic above the Néel temperature T_N when there is no long range magnetic order so that the unit cell is indeed that corresponding to the chemical formula. In fact, it is not. More quantitatively, one expects, according to the Slater picture, that the band gap E_g due to antiferromagnet order would be of order $k_B T_N$. It is actually forty times that! Thus the very existence of insulators (Mott insulators) with magnetic ions but without long range magnetic order (paramagnets) means that correlation effects are very important for such systems and that one-electron band theory does not correctly describe the electronic states.

An interesting feature of these insulators is the following. Since each ion has a magnetic moment, there are at each site, many degenerate states (say with different J_Z quantum numbers) only one of which is occupied. Despite there being a large number of energetically degenerate unoccupied states in the system, it requires a non-zero energy to transfer an electron from one site to another! An obvious question is whether a given substance could make a transition to such an insulating state from an uncorrelated metallic state. This is the Mott transition, which we now describe.

B. *Mott Transition*

In 1949, Mott[6] described a simple hypothetical model, namely a lattice of hydrogen atoms which can be expected to have a metal-insulator transition as a function of the lattice constant a. Clearly when $a \gg a_0$ (the Bohr radius), the system is best described as a collection of hydrogen atoms, and is an insulator with an activation energy $U \sim 12$ eV which is the sum of the energy I required to remove an electron from a hydrogen atom (ionization energy, $I = 13{\cdot}6$ eV) and to add it to another (electron affinity, $A = -1{\cdot}6$ eV). The activation or electron correlation energy $U = I + A$. However, for $a \ll a_0$, the electrons overlap strongly, an atomic description is not appropriate, and the system is a half-filled band metal. Clearly, as the single parameter a (lattice constant) is increased keeping the lattice structure unchanged, there is a metal to insulator transition as a increases to a value of order a_0. The two limits are described rather differently: For $a \ll a_0$, one is in the molecular orbital or Bloch regime whereas for $a \gg a_0$, one is in the Heitler London or atomic limit.

The obvious questions are:– Do such transitions occur? What are they like? As an experimental fact, they do not occur for a lattice of hydrogen atoms, which is insulating because the atoms form molecules (H_2) and the period doubled molecular crystal is a band insulator. Thus such effects could pre-empt the Mott transition. We now discuss a few candidate systems.

II. Some candidates for the Mott transition

A. *Phosphorus-Doped Silicon (Si:P)*

The donor electrons of phosphorus form a hydrogen-like state, bound to the P nucleus. The effective Bohr radius a_H^* is ~ 30 Å and the binding energy is $\sim 0{\cdot}03$ eV. Thus a collection of phosphorus atoms in silicon is like a collection of hydrogen atoms. Their density n can be increased by doping, and indeed at $n_c \sim 3{\cdot}7 \times 10^{18}$ P/cm^3 there is an insulator-metal transition. (This corresponds to $n^{1/3} a_H^* \sim 0{\cdot}25$, a dimensionless Mott criterion which works very well for a large number of metal insulator transitions).

This much-studied transition[7,8] is continuous. For example, the conductivity seems to go to zero continuously with an exponent ν equal to half ($\nu = 1/2$). The static dielectric ϵ constant

diverges on approaching metallic limit, with an exponent close to 2ν. The frequency and temperature-dependent conductivities go $\omega^{1/2}$ as $T^{1/2}$ and respectively. As one approaches the insulating phase from the metallic side, magnetic moments appear to form. For example, the magnetic susceptibility $\chi(T)$ is highly enhanced and strongly temperature-dependent; the spin fluctuations relax very slowly, the relaxation rate going to zero as $T^{1/2}$, in strong contrast to the linearity in temperature expected of a metal (Korringa relaxation rate). Recent NMR measurements show that even in the metallic state, a certain fraction of the donor electrons is localized; they have magnetic moments. In the insulating state, the low temperature magnetic susceptibility diverges as $T^{-1/2}$ close to the critical density. There is no long range (antiferromagnetic) order. All these observations point to the localization of electrons (no diffusion of charge or spin density) at the metal-insulator transition. The question is one of the operative mechanism. This is not settled as yet in spite of the wealth of experimental information, though there are many suggestions.

Mott argued[6,1] that because of long range coulomb forces, the transition will be discontinuous, from a state with no free carriers to a state with a non-zero density of them. In a lattice of hydrogen atoms, if an electron is removed from one site and added to another, the attraction between the proton left behind and the extra electron is coulombic; such a hole electron pair binds. However, if there are a large number of free carriers, the *screened* coulomb potential may be too weak to bind the electron hole pair; the electron hole excitation costs no energy and the system is metallic. Thus the carrier density is expected to increase discontinuously from zero (insulator) to a non-zero value (metal). Experimentally, the effective free carrier density changes continuously. In the absence of a detailed theory, it is not clear what is wrong with the argument.

A feature of the Si:P system which makes it different from a lattice of hydrogenic atoms is that it is not a periodic lattice, since the dopant phosphorus ions (protons of the hydrogenic atoms) go randomly into Si sites. Thus the system is intrinsically disordered. Electronic states in such a system can be localized if the randomness is large enough (Anderson localization)[9,10].

A system with Anderson-localized states at or near the Fermi energy is an insulator, irrespective of the average number of electrons per site (the Mott insulator requires one electron per site for an orbitally nondegenerate system; in general a commensurate number of electrons per site is needed). A perturbative scaling theory of the disordered electron gas has been developed by Finkelshtein[11], Lee, Kotliar and others[12] and the metal insulator transition in Si:P discussed in this light, with partial success. The perturbative theory neglects electronic commensuration effects. These are experimentally important. The conductivity exponent for the transition in Si:(P,B) [where the number of carriers ($n_P - n_B$) is less than the number of sites $n_P + n_B$] is close to unity rather the value half for Si:P. Also, the insulating state in Si:P has a measured gap against charge excitations, whereas this question is not addressed in the perturbative theory.

In brief, the Si:P system is a candidate for a correlation-driven metal insulator transition in a disordered system. The effects of correlation and disorder reinforce each other in ways, whose details are not yet clear.

B. V_{2-x} (Cr, Al)$_x$ O_3

This is another system extensively studied in the seventies because of the metal insulator transition in it. In the composition (x) temperature (T) plane the phase diagram is shown in figure 2[1,13]. There is a metallic phase M (with indications of spin density wave ordering in some regions of it), an antiferromagnetically ordered insulating phase AFI, and a paramagnetic insulating phase PI. The M-PI transition ends in a critical point, all other transitions being discontinuous. There is a structural change accompanying the metal-insulator transition. Because of this structural change and the effects of disorder (again unavoidable since substituents replace vanadium randomly), this system is not a perfect candidate for the Mott transition, though the general phase diagram is close to what is expected if increasing correlation is identified with increasing x (the Cr ions, being larger exert negative pressure). The discontinuous structural change could mean that electron lattice interactions are important in these systems. In the metallic phase, close to the transition, there are many indications of strong correlation effects (large T^2 term in resistivity, large enhancement of susceptibility and electronic specific heat). [Spalek[14] and co-workers have developed a theory of this metal insulator transition using approximate theories of correlated electrons.]

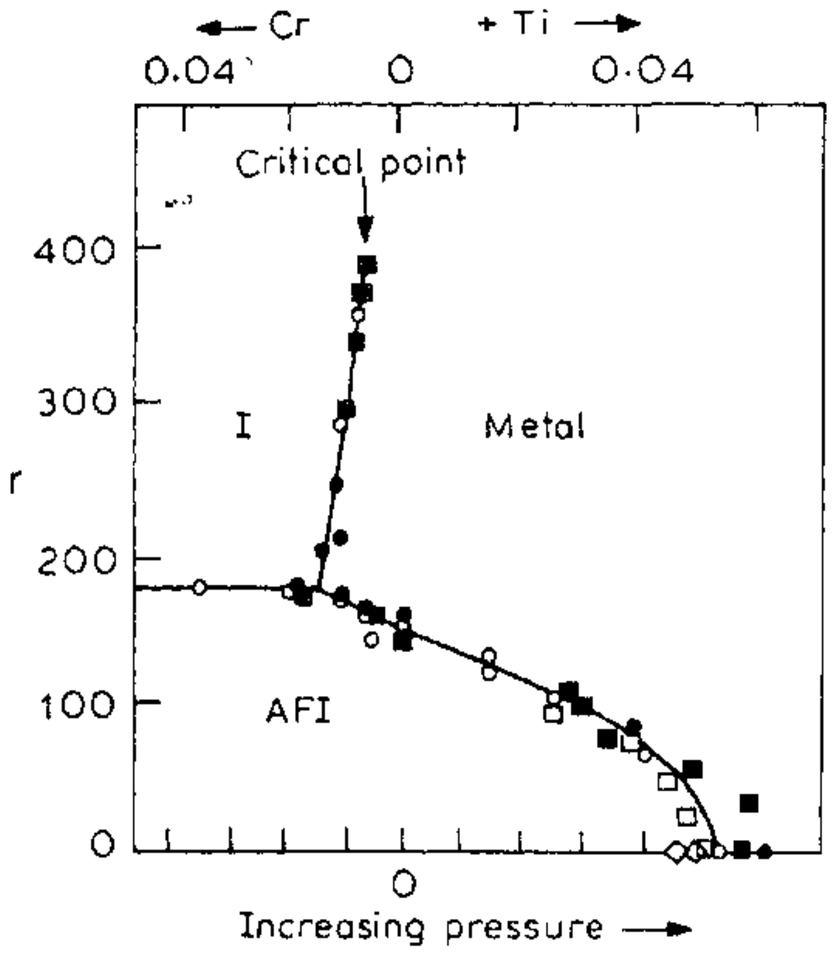

Figure 2. Metal-insulator transition in V_{2-x} (Cr, Ti)$_x$ O_3 where the amounts of Cr and Ti are as shown. The difference in sizes of Cr, Ti and V can be thought of as producing local (chemical) pressure. Cr is larger and exerts negative pressure while the smaller Ti exerts positive pressure. An equivalent pressure scale is shown on the x axis below. The phase diagram is shown in the temperature 'pressure' plane, temperature being in degrees Kelvin. Paramagnetic insulator (PI), metal (M) and antiferromagnetic insulator (AFI) phases are shown.

C. He^3

A rather different but cleaner system is He^3. Low temperature liquid He^3 is a pure Fermi system with strong, short range correlations between He^3 atoms. He^3 states are extended. In the solid phase, the fermion states are localized. Thus the liquid-solid transition can be considered (in this quantum Fermi system) to be analogous to the Mott metal-insulator transition. This system is free from possible complications due to 'frozen' disorder, electron lattice interactions and long range coulomb forces. It is however not a lattice system; the extended state regime is a liquid. But the liquid could be thought of as a cellular structure with a finite concentration of mobile atoms and no long range cellular order. Experimentally, the activation energy for He^3 migration in solid He^3 decreases to nearly zero as the fluid phase is approached[15]. Solid He^3 has a number of other unusual features[16], e.g. a specific heat linear in temperature at low temperatures. This is unknown in other crystalline insulating solids, where a T^3 dependence is obtained. The solid is extremely soft with respect to shear and has an unusual kind of magnetic ordering which requires long-range interactions. He^3 is well described as a nearly localized Fermi liquid[17]. The liquid solid or 'Mott' transition in He^3 is first order. But this could be because liquid-solid transitions are generically discontinuous.

D. La_2 Cu O_4 *etc.*

No Mott transition has been yet seen in this system, but La_2 Cu O_4 is an antiferromagnetic insulator with a Néel temperature of about 220K. It ought to have a partly filled band according to band theory and hence should have been a metal. It is therefore quite likely that the insulating antiferromagnetic state is due to correlation, i.e. that La_2 Cu O_4 is a Mott insulator[4]. La_{2-x} Sr_x Cu O_4 is then thought of as a Mott insulator with x holes per formula unit, i.e. a highly correlated system which is metallic because of a density of mobile holes. Thus understanding high temperature superconductivity probably requires that the Mott transition and the nature of the doped Mott insulator be elucidated.

It is almost certain, however, that in both NiO and La_2 Cu O_4, the lowest excitation energy involves transfer of a hole from the filled oxygen band to the Ni^{++} or Cu^{++} ion. This *d-p* charge

transfer energy is smaller than the energy difference U between Ni^{++}–Ni^{+++} or Cu^{++}–Cu^{+++} configurations. Are such charge transfer (d and p bands) insulators (which also require strong onsite d-d correlation for their existence) qualitatively different from one band (e.g. d band) Mott insulators? This is a problem of active interest for obvious reasons.

The above brief survey indicates that in experimental systems, the basic correlation-induced localization transition can be modified by effects due to long range interactions, electron lattice coupling, disorder, liquid-solid transition, electronic incommensuration, and low dimensionality. It is therefore important to first understand in detail the nature of the Mott transition and then to evaluate the effect of relevant perturbations such as those discussed. A simple model which can describe both strongly and weakly correlated systems as a parameter is varied and was first discussed in detail by Hubbard[18] nearly twentyfive years ago. This will concern us now.

III. Hubbard model

The Hubbard model describes electrons on a lattice, at sites i. The energy of an electron is ϵ. If two electrons (of opposite spin) are on the same site, their total energy is $(2\epsilon + U)$. U is thus a measure of repulsion between electrons. Electrons hop from a site i to a site j with an amplitude t_{ij}. Since hopping is due to overlap of wave functions, it is short-ranged and can be assumed to connect only the nearest neighbours j of a site i. The Hamiltonian can be written in terms of quantized field or creation and annihilation operators a, a^+ as follows:–

$$H = \sum_{i,\sigma} (\epsilon - \mu)\, a^+_{i\sigma}\, a_{i\sigma} + \sum_{ij\sigma} t_{ij}\, a^+_{i\sigma}\, a_{j\sigma} + U \sum n_{i\uparrow}\, n_{i\downarrow} \,. \tag{1}$$

The model has three dimensionless parameters, namely the following:–

(i) correlation energy U in units of the kinetic energy or bandwidth zt. i.e. $u = (U/zt)$,

(ii) electron density $1/N < \Sigma_{i,\sigma}\, n_{i\sigma} > = n$ or average number of electrons per site,

(iii) thermal energy $(k_B T/zt) = \theta$.

Two dimensionless parameters, namely u and θ (reduced correlation and reduced temperature) describe the half-filled case $n = 1$ which is the only density with the possibility of an insulating phase. The phase diagram in the space of these two variables is not known in detail, except for one dimension where for all u and $\theta = 0$ one has an insulator. The limits $u \ll 1$ and $u \gg 1$ are understood, but uncertainties are greatest for $u \sim 1$, the region of most interest for the Mott transition. The phases possibly depend on the dimensionality and on the lattice type as well. We describe below one-model system, and comment on other possibilities highlighting the uncertainties. We then discuss an approach which could provide detailed answers.

Consider a simple cubic lattice with nearest neighbour hopping only, and with one electron per site. The ground state is always antiferromagnetically-ordered and insulating. On increasing the temperature, there is a continuous transition from this ground state to a Pauli paramagnetic metallic state at small u and to a paramagnetic (local moment) insulator for large u (see figure 3). The small u system is a Slater insulator below the spin density wave temperature T_{SDW} and a metal above. The insulating state, or the gap against charge excitations is due to SDW order, disappearing with it. For large u, on the other hand, one has a Mott insulator: there is a gap against charge excitations in the paramagnetic phase, above the antiferromagnetic or Néel ordering temperature T_N. Thus, in the θ, u plane, above the magnetic ordering temperature there is a change from Pauli paramagnetic metal to paramagnetic insulator as u increases. This change is shown schematically by a dotted line in figure 3 and is the Mott transition. Is it a sharp transition, or a smooth crossover because of finite temperature effects? The basic point is that the condition for a charge gap to appear becomes different from that for antiferromagnetic order, somewhere in the region u, $\theta \sim 1$. How does this charge spin separation happen? What is the nature of charge and spin fluctuations in this regime? These questions are difficult to answer precisely because the Mott transition is between two qualitatively different regimes. We expand on this now.

In the small u regime, the antiferromagnetic order is due to a Fermi surface instability. The free-electron Fermi surface (for the nearest-neighbour hopping, exactly half-filled case) is the surface

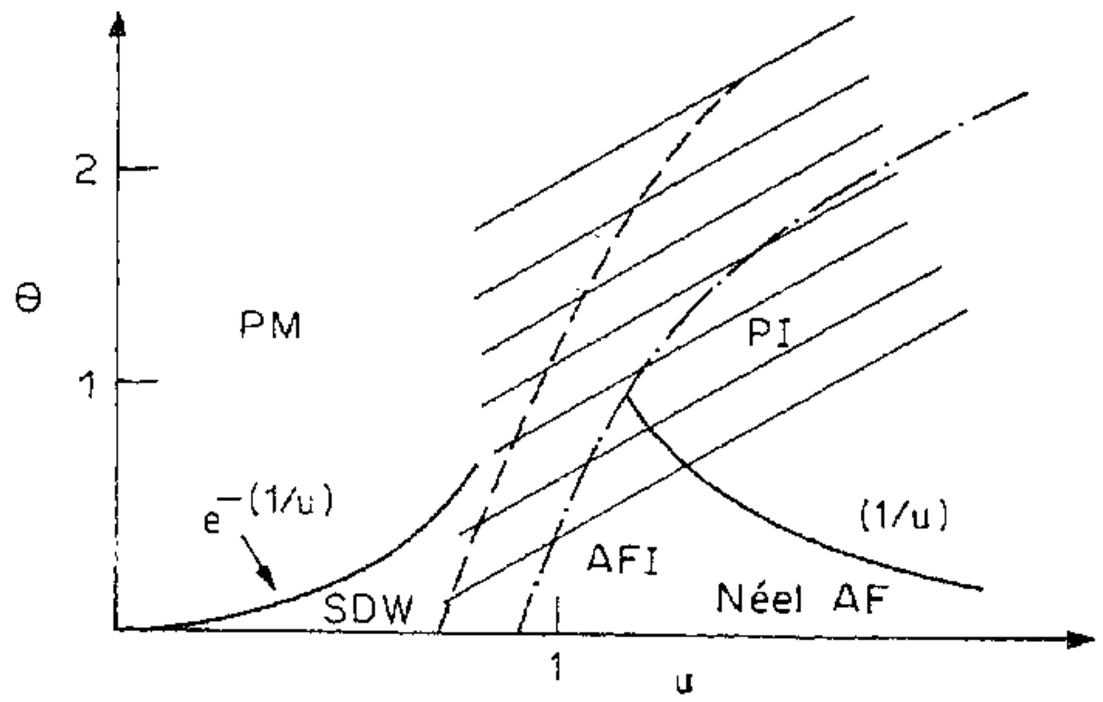

Figure 3. Schematic phase diagram of the half-filled Hubbard model in the reduced temperature (θ) and correlation (u) plane. The paramagnetic metal, paramagnetic insulator, and antiferromagnetic insulator regions are indicated. The shaded region is one of maximum uncertainty. Conjectured local moment formation boundary (dashed line) and the Mott transition line (dash-dotted line) are indicated.

of a cube and so consists of planes. The interaction u causes singular scattering of electrons from states near one planar segment to another parallel segment displaced by a wave vector $\mathbf{Q}$ ($|\mathbf{Q}| = (\pi/a)$ where a is the lattice constant). The spin susceptibility $\chi(\mathbf{Q})$ is logarithmically divergent at $T = 0$, and the ground state is a spin density wave with a commensurate wave vector $\mathbf{Q}$. The transition temperature θ_{SDW} is exponentially small i.e. $\theta_{SDW} \sim \exp(-1/u)$. The gap in the electron excitation spectrum, or the charge gap, is due to the SDW order and consequent new periodicity in the potential. The correlation is a singular perturbation on the mobile delocalized electrons.

For large u, the electrons are localized, one at each site and the hopping can be treated by perturbation theory. The most important effect, involving virtual charge fluctuations and of second order in t_{ij}, is an antiferromagnetic nearest-neighbour spin coupling $J_{ij} = (t_{ij}^2/(U)$. The antiferromagnetic Néel order thus occurs at a temperature $T_N \sim [(J_{ij})/k_B]$. The gap for charge excitations, known as the Mott-Hubbard gap, is of the order $U \gg k_B T_N$.

We notice that while the antiferromagnetic ordering is always to a two-sublattice arrangement, the transition temperature is not a monotonic function of u. There is no comprehensive theory which includes both limits. For example, if the weak correlation approach were extended to large u, T_{SDW} would appear to increase monotonically with u and to saturate at a value $\theta \sim 1$ for $u \to \infty$. Clearly, this cannot be taken seriously since u cannot be treated perturbatively when large. On the other hand, if the correlation dominated or large u picture is used, hopping is treated as a perturbation and this would suggest that T_N increases as $(1/u)$ increases, or as u decreases. The basic localized electron picture is expected to break down, though, for $u < 1$.

The change from SDW order to Néel order at low temperatures with increasing u, is of course a reflection of the Mott metal-insulator transition occurring at a higher temperature. If this $T \neq 0$ transition is continuous i.e. the Mott-Hubbard gap evolves continuously it may not show up as a sharp change in electrical conductivity since at any non-zero temperature, the insulating state also has an activated non-zero conductivity. Like any continuous metal-insulator transition (e.g. Anderson localization), the Mott transition will then probably show up most clearly if it occurs at $T = 0$. For this to happen, the low temperature magnetic ordering should be circumvented. This can be done in model systems by a choice of lattice or of suitable further neighbour hopping. Another possibility is that the system is random (e.g. Si:P). In this case again, antiferromagnetic long-range order is
here may be a continuous zero temperature conductivity and excitation gap transition
critical u. However, the transition could be strongly affected in character by disorder,
to localize electrons. The extent to which the $T = 0$ fixed point affects observed
is an obvious question of interest. Another possibility is that the Mott transition
her in the short-range correlation model described above or in real systems

because of other factors such as electron lattice coupling (V_2O_3 ?), coulomb interaction effects, or solidification (He^3 ?). Such a discontinuous transition line may terminate in a critical point somewhere in the (u, θ) plane (doped V_2O_3 ?).

In the absence of a reliable detailed theory, the region of the Hubbard model phase diagram, shown shaded (figure 3) is not well-charted; some possibilities have been indicated above. We conclude by briefly mentioning earlier theoretical work, and our ideas.

IV. Theories of the Hubbard model

Earlier attempts at understanding the Hubbard model have been reviewed by many authors (Ref. 3 is one such). We shall consider here two sets of ideas, one based on correlated wave functions due to Gutzwiller, and the other on local moment formation and the 'alloy analogy'.

The effect of onsite repulsion U is to energetically disfavour states with double occupancy ($_\uparrow$ and $_\downarrow$ spin states both occupied) at any given site. A wavefunction incorporating this effect is $|\psi\rangle = \{\Pi_i [1 - g n_{i\uparrow} n_{i\downarrow}]\}| \psi_0\rangle$ where $|\psi_0\rangle$ is a suitable independent fermion or 'Hartree Fock' many body wavefunction. For $g = 0$, one has no correlation, and for $g = 1$, no double occupancy. One can calculate g as a function of u by minimizing the energy. This cannot be done exactly, but an approximate estimate for the simple cubic lattice gives $u_0 = 4$. For $u > u_0$ one has a Mott insulating state since each site has exactly one electron, and for $u < u_0$ a metallic state since the ground state has charge fluctuations (which need double occupancy). One is thus tempted to think of g, or double occupancy, as a kind of order parameter for a continuous insulator-metal transition since the latter continuously evolves from zero as u decreases from u_0. This cannot be quite true for, from simple large u perturbation theory the average number of doubly-occupied sites is a fraction of order $(1/u)^2$. In a metal, the doubly-occupied sites and holes are free, whereas in the insulator they are bound so that charge fluctuations have a gap. The Gutzwiller description, while qualitatively correct for the strongly correlated metal close to the Mott transition ($u \lesssim u_0$) has an infinite Mott-Hubbard gap in the insulating phase, even close to u_0. The Gutzwiller projection is a very direct and useful procedure however for constructing approximate wavefunctions for strongly-correlated lattice Fermi systems, especially if $|\psi_0\rangle$ is chosen to reflect other correlations present in them.

The much older approach based on local moment formation and the 'alloy' analogy[19,20,21] proceeds as follows:- As u increases each site is singly occupied with increasing probability, and thus a local magnetic moment appears there beyond critical value u_c (see figure 3 for a schematic diagram). This condition can be formulated as the divergence of the on-site spin or magnetic susceptibility χ_{ii}. As u increases further, the average local moment at each site becomes larger. Assuming that the size of the moment and just two directions (up and down) characterize it, i.e. assuming an Ising magnetic moment, the system can be thought of as a two-component 'alloy'. The up and down pointing moments are the two constituents. The density of single-particle states for such a binary alloy can be approximately calculated, e.g. in a single site approximation such as the CPA. When the magnitude of the local moment exceeds a certain value u_M (see figure 3 for a schematic indication) there is a gap in the density of states, i.e. in the spectrum of single particle excitations. The Mott-Hubbard gap grows to a value of order U as u increases. The interaction between magnetic moments at neighbouring sites can also be obtained approximately. A phase diagram similar to that of figure 3 but with details in the shaded region based on rather arbitrary approximations and assumptions has been obtained, for example, by Economou and co-workers[20,21].

In the 'alloy' analogy, one clearly breaks rotational invariance at each site. One replaces moments fluctuating in direction and in time by static moments oriented up or down and thus approximates a translationally-invariant system by an alloy. The analogy however emphasizes the single-site nature of the Mott-Hubbard gap and is the only approximation to provide a reasonable description of it. We now describe a reformulation and development of these ideas in a language which seems more promising. One way of considering the moment problem is to describe the free energy of the Hubbard model in terms of a vector moment $\mathbf{m}_i(\tau)$ at each site i; in general $\mathbf{m}_i$ depends on time τ also. The free energy functional $F\{\mathbf{m}_i(\tau)\}$ can be approximately calculated using a Hubbard-Stratonovich transformation. It has, e.g. for the static functional $F\{\mathbf{m}_i(z_m = 0)\} = F\{\mathbf{m}_i\}$ a part F_l which depends only on the magnitude of $\mathbf{m}_i$'s, i.e. on $|\mathbf{m}_i|^2$ and another F_p which depends on the angle between them, e.g. on $(\mathbf{m}_i \cdot \mathbf{m}_j)$. If F_l has a minimum for non-zero $|m_i|^2$ one has local

moments. Their interaction, as described by F_p determines the long-range magnetic order at low temperatures. The one-electron density of states for large u depends essentially on the magnitude of the local moment and not on its direction, i.e. it is first calculated assuming only F_l to be non-zero and is expected to develop a gap for a large enough u. If the effect of moment interactions (without long range order) and of temporal fluctuations can be included and shown not to destroy this Mott-Hubbard gap, one has a complete theory for the strongly correlated regime. The appearance of the density of states gap is the Mott transition. As discussed by several authors, the insulating phase has a local gauge invariance. This important property as well as the entire scenario outlined are best discussed using a particle hole transformation first considered by Kemeny[22], which we develop below.

The up-spin electron is considered one species of fermion (say 1) and the down-spin hole is considered another species of fermion. Thus one writes $a^{\dagger}_{i\uparrow} = a^{\dagger}_{i1}$ and $a^{\dagger}_{i\downarrow} = a^{\dagger}_{i2}$ so that the Hubbard-Hamiltonian can be written as

$$H = \sum_{i,j} t_{ij} (a^{\dagger}_{i1} a_{j1} - a^{\dagger}_{i2} a_{j2}) - U \sum_{i} n_{i1} n_{i2} + (U/2 - \mu) \sum_{i} n_{i1} - n_{i2}) \; . \tag{2}$$

The two fermions thus attract each other locally, with a strength U. Further, in the half-filled case, one has $\langle n_{i1} \rangle = \langle n_{i2} \rangle$. This clearly means from equation (2) that the corresponding 'magnetic field' $(U/2 - \mu) = 0$, or that $\mu = U/2$. Deviation from half-filling is described by the chemical potential μ moving away from $(U/2)$.

Because of the local attraction between Fermions 1 and 2 a 12-pair is possible. This pairing is just a local moment formation. The local-order parameter $\langle a^{\dagger}_{1i} a^{\dagger}_{2i} \rangle$ corresponds to a moment in the transverse direction (the operator $a^{\dagger}_{1i} a^{\dagger}_{2i}$ is the same as S^{+}_{i}). The triad of parameters $\langle a^{\dagger}_{1i} a^{\dagger}_{2i} \rangle$, $\langle a_{1i} a_{2i} \rangle$ and $\langle n^{-}_{1i} n_{2i} - 1 \rangle$ are the components S^{+}_{i}, S^{-}_{i} and S^{z} of the vector $\mathbf{S}_i$ or $\mathbf{m}_i$. Starting from the Hamiltonian equation (2) one can find free energy as a functional of $\{\mathbf{m}_i\}$.

A few low-order terms in a polynomial expansion for $F\{m_i\}$ are shown in figure 4. They can be unambiguously divided into terms that are independent of the relative phase angle of the moments at i and j and those that depend on it. The second-order term F^{p2}_{ij} is long-ranged in space and its Fourier transform coefficient $\chi^{-1}(\mathbf{Q})$ has a logarithmic singularity for $|\mathbf{Q}| = \pi/a$ at $T = 0$. This indicates an SDW instability, without local moment formation. As u increases, the single site or local terms are more important since u is site local. This has several consequences. For example, the free energy minimum is for $|m^{2}_{i}| \neq 0$ (figure 5). Further, in the term F^{p}_{ij} as u increases, all local processes occurring at sites i and j must be summed over first, and the effective interaction energy couples moment vertices $\{\delta F^{l}/\delta m_i\}$ and $\{\delta F^{l}/\delta m_j\}$. (see figure 6). These are proportional to $\mathbf{m}_i$ and $\mathbf{m}_j$

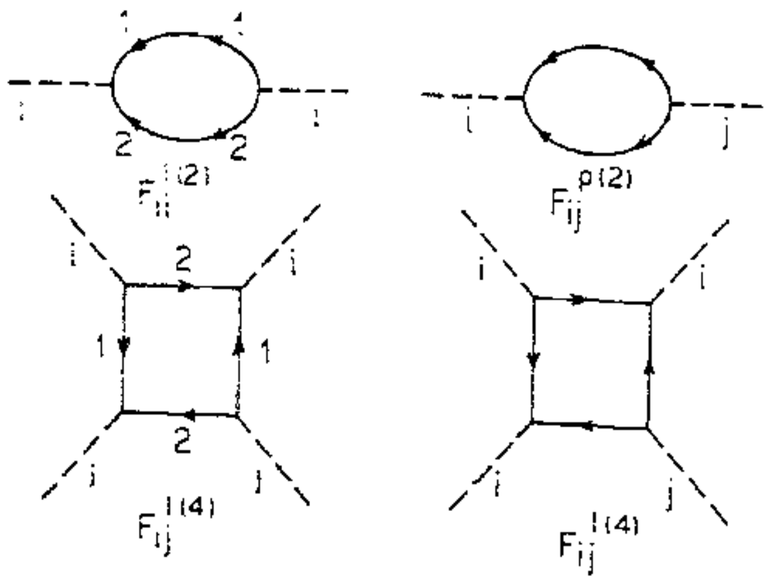

Figure 4. Some low-order diagrams for the free energy as a functional of the local moment $\mathbf{m}_i$. The local terms F^{l} do not depend on the angle between moments $\mathbf{m}_i$ and $\mathbf{m}_j$, whereas the terms F^{p} do.

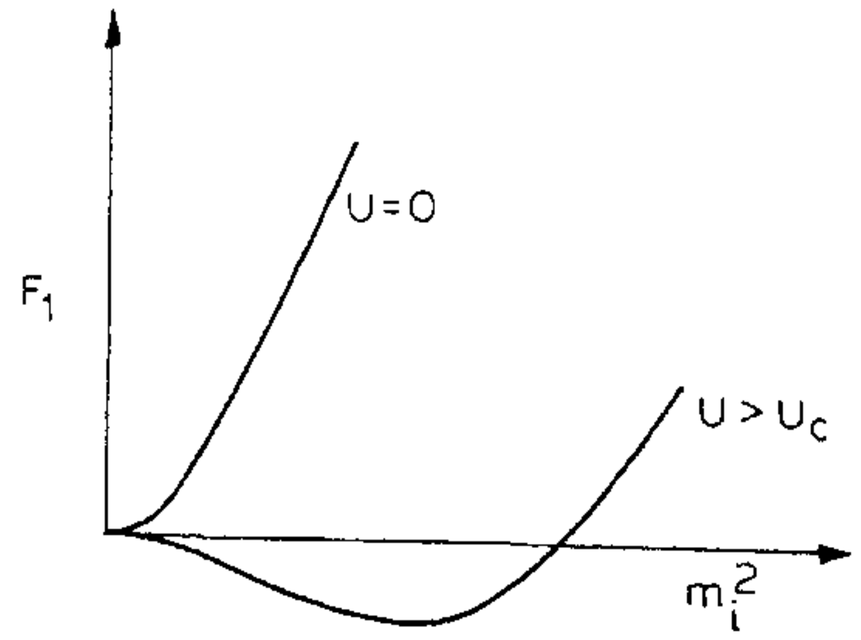

Figure 5. Local part of the free energy as a function of the size m^{2}_{i} of the local moment, for two cases, namely $U = 0$ and $U > U_c$ (schematic).

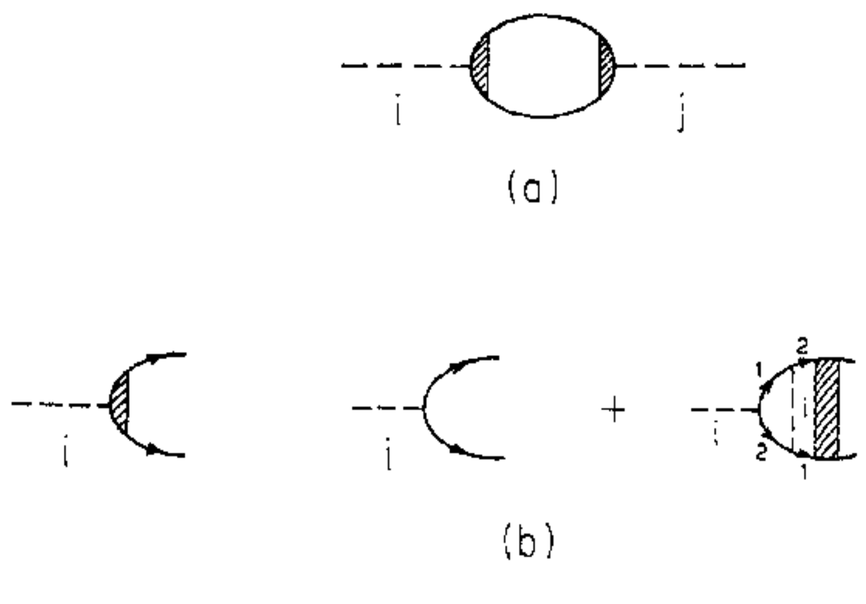

Figure 6. Effective coupling between moments at sites i and j.

respectively. Thus the effective interaction energy has the Heisenberg form $\mathbf{m}_i \mathbf{m}_i J_{ij}$ where J_{ij} can be calculated from the renormalized propagators $G_{ij} G_{ji}$ and the vertices to be (t_{ij}^2/U) for large u. Thus one can see how, with increasing u, magnetic interactions change from long-range, small, magnetic fluctuations in the electron gas to short-range coupling between well defined, large, local moments.

As u increases, the one-electron propagator also is dominated by single-site processes. The strong on-site potential due to a local moment binds the electron; the process is schematically shown in figure 7. A consistent approximation for the electron self-energy due to these processes is to consider the diagrams for F^I $(\{\mathbf{m}_i\})$ and functionally differentiate this with respect to the local electron number operator $a_{i\sigma}^{\dagger} a_{i\sigma}$. This binding is the Mott-Hubbard gap which increases from zero as u grows beyond $u_m > u_c$.

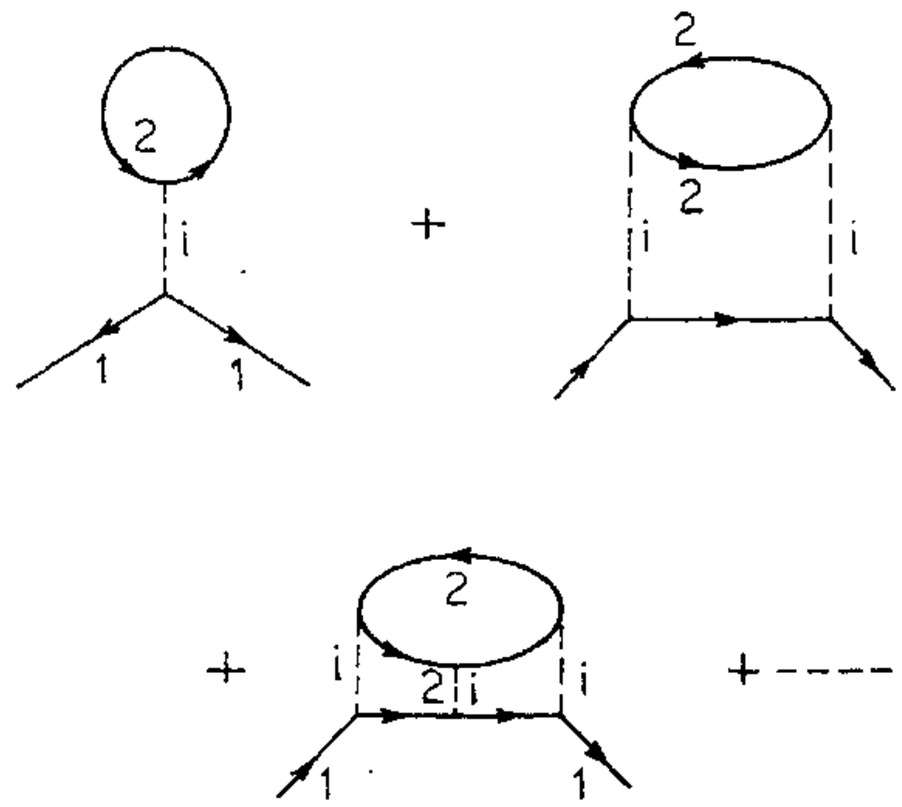

Figure 7. Interaction between an electron and the local moment at site i.

The diagrammatic summations indicated above can be carried out systematically in a $(1/z)$ expansion, as will be shown in a future publication[23]. We have outlined here how such a programme makes explicit the sequence of approximations involved, and the dominance of local correlations as u increases.

One can thus develop a fairly detailed and explicit picture of local moment formation due to correlations, and the formation of a Mott-Hubbard gap (as a bound state) when the moments are large. Both of these are single site effects and have to be described properly first when u is large. The residual intersite effects depending on the angle between moments $\mathbf{m}_i$ and $\mathbf{m}_j$ are relatively small perturbations for large u. Away from half-filling, the number of 1 and 2 fermions is not the same, and there are always some unpaired fermions, i.e. sites without moments. The system thus consists of moving, fluctuating moments and holes. A detailed theory of this regime is clearly of great possible interest for doped oxide metals and superconductors.

V. Conclusion

We have attempted here to describe qualitatively the Mott transition and a few systems where this might occur. Experimentally, most candidate systems are not clean realizations of the proposed continuous transition. This could be because the continuous transition is masked by relevant perturbations such as disorder electron lattice coupling, coulomb interactions etc. Therefore the simplest short-range correlation model likely to show a continuous Mott transition, namely the Hubbard model needs to be understood in the transition region and effects of various perturbations need to be analysed. We do not have such an understanding yet; some attempts have been described. It is argued that a many-body approach focusing on the development of local moments and 'bound' states due to them, has the potential to provide a detailed description of the metal-insulator transition and related systems such as the strongly correlated metal.

Note added in proof: Between the Raman Centenary Symposium in December 1988 the publication of its proceedings, there have been many significant advances in our understanding of correlated electronic systems and of the Mott transition. In particular, H. R. Krishnamurthy and coworkers have developed [*Phys. Rev. Lett.* **64,** 950 (1990)] a bosonic mean field theory for the $T = 0$ Mott-Hubbard transition in a triangular two-dimensional lattice, and have pointed out that inclusion of fluctuation effects is likely to describe the *entire* phase diagram shown in Figure 2 of this paper.

References

1. See *Metal Insulator Transitions,* by N. F. Mott (Taylor and Francis, London 1974) for a comprehensive review of phenomena and theories, as well as early references.
2. de Boer, J. H. and Verwey, E. J. W., *Proc. Phys. Soc.*, 1937, **A49,** 59.
3. Ramakrishnan, T. V., The Metal-Nonmetal Transition in *Metallic and Nonmetallic States of Matter,* (eds. Edwards, P. P. and Rao, C. N. R.), Taylor and Francis, London, 1985.
4. Anderson, P. W., Fifty Years of the Mott Phenomenon: Insulators, Magnets, Solids and Superconductors as Aspects of Strong Repulsion Theory in *Frontiers and Borderline in Many Particle Physics,* (eds. Schrieffer, J. R. and Broglia, R. A.) North Holland, Amsterdam, 1987.
5. Slater, J. C., *Phys. Rev.*, 1951, **59,** 537.
6. Mott, N. F., *Proc. Phys. Soc.*, 1949, **A62,** 416.
7. Milligan, R. F., Rosenbaum, T. F., Bhatt, R. N. and Thomas, G. A. in *Electron Electron Interactions in Disordered Systems* (eds. Efros, A. L. and Pollak, M.) North Holland, Amsterdam, 1985.
8. See articles in *Metal Insulator Transitions,* (ed. Finlayson, D. M.) Proc. 31st Scottish Universities Summer School, Edinburgh University Press, Edinburgh, 1987.
9. Anderson, P. W., *Phys. Rev.* 1958, **109,** 1492.
10. See Lee, P. A. and Ramakrishnan, T. V., *Rev. Mod. Phys.*, 1985, **57,** 287.
    Ramakrishnan, T. V., in *Chance and Matter,* (eds. Souletie, J., Vannimenus, J., Stora, R.) Elsevier, Amsterdam, 1987.
11. Finkelshtein, A. M., *Sov. Phys., JETP,* 1983, **57,** 97, *J. Phys.* 1984, **B56,** 189.
12. See for example Castellani, C., Kotliar, G. and Lee, P. A., *Phys. Rev. Lett.*, 1987, **59,** 323.
    Kirkpatrick, T. R. and Belitz, D., 1989, Univ. of Maryland preprint.
13. Mcwhan, D. B. and Remeika, J. P., *Phys. Rev.* 1970, **B2,** 3734.
14. Spalek, J., Datta, A. and Honig, J. M., *Phys. Rev. Lett.*, 1987, **59,** 728.
15. See, e.g. Kumar, P. and Sullivan, N. S., *Phys. Rev.*, 1987, **B35,** 3162.
16. See, e.g. *The Physics of Liquid and Solid Helium,* Vols. I and II, (eds. K. H. Bennmann and J. B. Ketterson) Wiley Interscience, New York, 1976 – 78.
17. Vollhardt, D., *Rev. Mod. Phys.*, 1984, **56,** 99.
18. Hubbard, J., *Proc. R. Soc.*, 1963, **A276,** 238.
19. Hubbard, J., *Proc. R. Soc.*, 1964, **A281,** 401.
20. Cyrot, M., *J. Phys.* (Pairs), 1972, **33,** 125.
21. Economou, E. N., White, C. T. and DeMarco, R. R., *Phys. Rev.* 1978, **B18,** 3946, 3959, 3968.
22. Kemeny, G. and Caron, L. G., *Rev. Mod. Phys.*, 1968, **40,** 790.
23. Ramakrishnan, T. V., to be published.