ARTICLES

Charge-transfer metal-insulator transitions in the spin-$\frac{1}{2}$ Falicov-Kimball model

Woonki Chung and J. K. Freericks

Department of Physics, Georgetown University, Washington, DC 20057-0995

The spin-$\frac{1}{2}$ Falicov-Kimball model is solved exactly on an infinite-coordination-number Bethe lattice in the thermodynamic limit. This model is a paradigm for a charge-transfer metal-insulator transition, where the occupancy of localized and delocalized electronic orbitals rapidly changes at the metal-insulator transition (rather than the character of the electronic states changing from insulating to metallic as in a Mott-Hubbard transition). The exact solution displays both continuous and discontinuous (first-order) transitions. [S0163-1829(98)02419-9]

I. INTRODUCTION

The Falicov-Kimball model (FKM) (Ref. 1) is the simplest model for a charge-transfer metal-insulator transition. In this model there are two types of electronic states: (i) localized f or d orbitals which have a negligible overlap with neighboring electronic orbitals but possess a strong on-site Coulomb repulsion (usually taken to be infinite), and (ii) delocalized conduction-band orbitals in which the Coulomb repulsion between two conduction electrons is neglected. The only ''dynamic'' Coulomb interaction that is included is the Coulomb repulsion between a conduction electron and a localized electron that occupy the same lattice site. The metal-insulator transition takes place when there is a rapid change in the thermodynamic occupation of the electron levels as a function of temperature (or some other thermodynamic variable such as pressure). Hence an insulator (or semiconductor), which has most electrons lying in the localized states, rapidly changes its character to a metal as the electronic charge transfers from the localized levels to the conduction band. This transition is similar to the liquid-gas phase transition, in which the density changes abruptly at the first-order transition temperature, or smoothly varies when the temperature is above the critical point. Here it is the density of electrons in the conduction band that changes at the transition.

It was originally thought that this model described the most important physics behind the metal-insulator transitions in a variety of transition-metal and rare-earth compounds. These materials (such as SmB_6, V_2O_3, Ti_2O_3, NiS, etc.) exhibit a variety of metal-insulator transitions with either continuous or discontinuous changes in the conductivity occurring as the temperature is varied. This opinion, however, was not shared by everyone, since most of these materials simultaneously exhibit structural and metal-insulator phase transitions, and because the FKM ignores all effects arising from the hybridization between the localized and conduction-band orbitals (such as the Kondo effect, and the screening of the localized moments). It remained unclear whether it was the electronic system, modeled by the FKM,

that was driving the transition, or whether it was driven by other (electron-phonon or hybridization) effects.

Furthermore, there are competing scenarios for metal-insulator transitions. In the Mott-Hubbard scenario,[2] it is the electron-electron correlations that change *the character of the electrons* (within a single band) from an insulator to a conductor. In the Anderson picture,[3] however, it is disorder that produces states that are localized and delocalized (within a single band), and the metal-insulator transition takes place by adjusting the Fermi level between the localized and delocalized states.

Recently, however, a series of experiments[4] have been performed on NiI_2, which appears to be a paradigm for a charge-transfer metal-insulator transition as described by the FKM. In this material, the Ni ions donate one electron to each of the neighboring I ions, filling the iodine p shell. The Fermi level lies within the d band of the Ni ions. This material is an insulator because of the strong electron correlations within the d band (which also lead to antiferromagnetic order at low temperatures). As the pressure is increased, the Néel temperature increases, and then disappears at the same point where the conductivity has a discontinuous increase by two orders of magnitude. Detailed x-ray-diffraction studies showed that there was no structural phase transition occurring at the metal-insulator transition. Instead, the experimental evidence points toward a transfer of charge from the I ions to the Ni ions, which quench the local moments (changing Ni^{2+} to Ni^+), and leave behind conduction holes in the iodine p band. Such a scenario is described by a FKM.[5]

The theoretical side of the FKM has also been controversial. The original solution of the model[1] involved just a mean-field analysis of the interband Coulomb interaction. That approximate solution found both continuous and discontinuous metal-insulator transitions. Later, an approximate version of the coherent potential approximation (CPA) was applied to the model by different groups, and produced conflicting results:[6,7] the first-order metal-insulator transitions were obtained in one solution;[6] while only continuous transitions were obtained in the other solution.[7] Hence it is important to be able to solve the FKM in an approximation-free

fashion (and in the thermodynamic limit), to resolve the question of whether or not it contains discontinuous metal-insulator transitions.

We present an exact solution of the simplest spin-$\frac{1}{2}$ Falicov-Kimball model on an infinite-coordination Bethe lattice (in the thermodynamic limit). We find that the solution possesses both continuous and discontinuous metal-insulator transitions. In Sec. II, we describe the formalism used in solving the problem, and describe the numerical techniques employed. In Sec. III, we present our results, and end with our conclusions in Sec. IV.

II. FORMALISM

The FKM consists of two types of electrons: a localized (dispersionless) valence band and a conduction band (separated by an energy gap Δ). The Fermi level lies within the energy gap at zero temperature, so that all of the electrons *a priori* occupy the localized valence band, and the system is an insulator. Hence the FKM was originally described within an electron-hole picture, where one considers holes within the valence band. The valence holes have a direct Coulomb repulsion $U_{\widetilde{f}\widetilde{f}}$, which disfavors two holes occupying the same localized orbital. In addition, there is an interband on-site electron-hole Coulomb interaction ($-U<0$) which drives the charge-transfer metal-insulator transitions. The Coulomb interaction is attractive, because the electron and hole have opposite charge. The resulting Hamiltonian is

$$H=\sum_{\mathbf{k},\sigma} \left[\epsilon(\mathbf{k})+\Delta+\frac{W}{2}\right]d^{\dagger}_{\mathbf{k},\sigma}d_{\mathbf{k},\sigma}-U\sum_{i,\sigma\sigma'} d^{\dagger}_{i\sigma}d_{i\sigma}\widetilde{f}^{\dagger}_{i\sigma'}\widetilde{f}_{i\sigma'}$$

$$+U_{\widetilde{f}\widetilde{f}}\sum_{i} \widetilde{f}^{\dagger}_{i\uparrow}\widetilde{f}_{i\uparrow}\widetilde{f}^{\dagger}_{i\downarrow}\widetilde{f}_{i\downarrow}, \tag{1}$$

where $d^{\dagger}_{\mathbf{k},\sigma}$ ($d_{\mathbf{k},\sigma}$) is the creation (annihilation) operator for a conduction electron of wave vector $\mathbf{k}$ and spin σ, $\epsilon(\mathbf{k})$ is the dispersion of the conduction band (with bandwidth W), and $\widetilde{f}^{\dagger}_{i\sigma}$ ($\widetilde{f}_{i\sigma}$) is the creation (annihilation) operator for a localized hole at lattice site i (the on-site energy of the localized hole is chosen as the origin of the energy axis). Since $U_{\widetilde{f}\widetilde{f}}$ is large[1] for most localized levels, we choose the limit where $U_{\widetilde{f}\widetilde{f}}\rightarrow\infty$, and restrict the number of f holes per site to $n_{\widetilde{f}}\leqslant 1$. Since the d electrons in the conduction band originate by thermal excitation from the f band, their number is constrained to $n_d=n_{\widetilde{f}}$ in order to conserve the total number of electrons.

We have found it to be more convenient for our numerical calculations to study the model in the particle picture rather than in the original hole picture given in Eq. (1). We employ a particle-hole transformation ($\widetilde{f}_{i\sigma}\rightarrow f^{\dagger}_{i\sigma}$ and $\widetilde{f}^{\dagger}_{i\sigma}\widetilde{f}_{i\sigma}\rightarrow 1-f^{\dagger}_{i\sigma}f_{i\sigma}$), and represent the kinetic energy in the localized basis to transform the Hamiltonian into

$$H=-\sum_{ij,\sigma} t_{ij}d^{\dagger}_{i\sigma}d_{j\sigma}+E_f\sum_{i,\sigma} f^{\dagger}_{i\sigma}f_{i\sigma}-\mu\sum_{i,\sigma} (d^{\dagger}_{i\sigma}d_{i\sigma}+f^{\dagger}_{i\sigma}f_{i\sigma})$$

$$+U \sum_{i,\sigma\sigma'} d^{\dagger}_{i\sigma}d_{i\sigma}f^{\dagger}_{i\sigma'}f_{i\sigma'}+U_{ff}\sum_{i} f^{\dagger}_{i\uparrow}f_{i\uparrow}f^{\dagger}_{i\downarrow}f_{i\downarrow}, \tag{2}$$

where $d^{\dagger}_{i\sigma}$ ($d_{i\sigma}$) is the creation (annihilation) operator for a conduction-band electron of spin σ at site i, t_{ij} is the hopping matrix between lattice sites i and j [which yields the band structure $\epsilon(\mathbf{k})$], $E_f=U-\Delta-W/2$ is the localized electron site energy measured from the middle of the conduction band, and U_{ff} is the on-site Coulomb repulsion between f electrons. A chemical potential μ is introduced to satisfy the constraint $n_d+n_f=1$ as $U_{ff}\rightarrow\infty$. This Hamiltonian has also been used as a model for intermediate-valence problems,[8] as a thermodynamic model for an annealed binary alloy,[9] as a simplified Hubbard model,[10] and as a model for metamagnetism and anomalous magnetic response[11] (when a magnetic field is added).

We choose to solve the FKM in the infinite-coordination-number limit following the method of exact solution developed by Brandt and Mielsch[12] and expanded by Freericks and Zlatić.[11] In the infinite-coordination-number limit, the local approximation becomes exact, implying that one can neglect the momentum dependence of the irreducible self-energy and the irreducible vertex functions, but one needs to determine the frequency dependence. Hence the problem on the lattice can be mapped onto a problem on a single site, but coupled to an effective medium, which represents the dynamical information of all of the other sites of the lattice. The effective medium needs to be determined self-consistently in order to solve the lattice problem exactly.

We begin with the local Green's function of the conduction electrons for each spin, $G(i\omega_n)\equiv G_n$ evaluated at the fermionic Matsubara frequencies $\omega_n=(2n+1)\pi k_B T$, and express it explicitly in terms of the "bare" Green's function $G_0(i\omega_n)$ (which contains all of the dynamical information of the other sites in the lattice):

$$G_n=\frac{1-n_f}{G_0^{-1}(i\omega_n)}+\frac{n_f}{G_0^{-1}(i\omega_n)-U}, \tag{3}$$

where

$$n_f=\left[1+q\,\exp\{(E_f-\mu-U)/k_B T\}\right.$$

$$\left.\times\prod_n \{1-UG_0(i\omega_n)\}^{-1}\right]^{-1}. \tag{4}$$

Here $q=(2J_0+1)/(2J+1)$ is the ratio of the spin degeneracies without (J_0) and with (J) an f electron at a lattice site, and we take $q=\frac{1}{2}$. (This is a different value from the original FKM work, which took $q=2$. We do not expect the results to depend strongly on the value of q.) The "bare" Green's function also satisfies the usual Dyson equation

$$G_0^{-1}(i\omega_n)=G_n^{-1}+\Sigma_n, \tag{5}$$

which can be viewed as a definition of the self-energy $\Sigma(i\omega_n)\equiv\Sigma_n$. The loop for determining the Green's functions is completed by evaluating the self-consistency equation by summing the momentum-dependent Green's function over all momenta [i.e., integrating over the noninteracting density of states $D(\epsilon)$],

$$G_n=\int_{-\infty}^{\infty}d\epsilon\frac{D(\epsilon)}{i\omega_n+\mu-\epsilon(\mathbf{k})-\Sigma_n}, \tag{6}$$

yielding an exact solution of the model.

The algorithm for solving the FKM is the same as that used for numerically solving the Hubbard model:[13] (i) Begin with an initial self-energy Σ_n(init.) [we chose either Σ_n(init.)$=0$, or we interpolated from a higher-temperature run]; (ii) use Eq. (6) to find G_n; (iii) then use Eq. (5) to find $G_0(i\omega_n)$. (iv) Next determine n_f from Eq. (4), and (v) determine a new G_n from Eq. (3). (vi) Finally, use Eq. (5) with the new Green's function and the old "bare" Green's function to extract a new self-energy, and (vii) go back to step (ii) to repeat the iteration until convergence is reached.

Since we are interested in dynamical properties, we also need to solve for the retarded Green's functions on the real axis. We do this by first performing an imaginary-axis calculation to find the filling of the f electrons (n_f), and then solving the analytically continued equations (3), (5), and (6) where the Matsubara frequencies are simply replaced by the real frequencies ($i\omega_n \rightarrow \omega + i\delta$). Convergence of these equations is rapid under iteration on both the real and imaginary axes (less than 100 iterations on average for a convergence of the self energy to one part in 10^9 on the imaginary axis, and to one part in 10^3 on the real axis).

Note that the self-consistent equations for fixed n_f, i.e., Eqs. (3), (5), and (6), are identical to those employed in the CPA. However, n_f is explicitly determined by Eq. (4) in the infinite-coordination-number limit, while, in the CPA, it is determined thermodynamically by minimizing the trial free energy[6,7] $F_{CPA}[n_f]$ as a function of n_f:

$$F_{CPA}[n_f] = -2k_B T \int_{-\infty}^{\infty} d\epsilon \, A(\epsilon) \, \ln(1 + e^{-\epsilon/k_B T})$$
$$+ \left(\mu - U + \Delta + \frac{W}{2} \right)(1 - n_f)$$
$$+ k_B T \left[n_f \ln n_f + (1 - n_f) \ln \frac{1 - n_f}{q} \right], \quad (7)$$

which, for the conduction electrons, has the form of a noninteracting electron system with the noninteracting density of states $D(\epsilon)$ replaced by the interacting one, $A(\epsilon) = (-1/\pi) \text{Im} G(\epsilon)$. Here ϵ is measured from the chemical potential μ. Since the conduction electrons are effectively noninteracting for fixed value of n_f (Ref. 14) (i.e., the interacting density of states for fixed n_f does not vary with temperature), the trial free energy in Eq. (7) is also exact. Naturally, the n_f obtained in the CPA by minimizing the trial free energy agrees with the n_f found within the Brandt-Mielsch formalism using Eq. (4).

We solve the model on an infinite-coordination-number Bethe lattice. As the coordination number Z increases, the hopping integral t between the nearest-neighbor sites is scaled as $t \rightarrow t^*/\sqrt{Z}$ in order to have a nontrivial kinetic energy, and the density of states for the noninteracting system becomes Wigner's semicircle with $W = 4t^*$:

$$D(\epsilon) = \frac{1}{2\pi t^{*2}} \sqrt{4t^{*2} - \epsilon^2}. \quad (8)$$

We take t^* as our energy unit ($t^* = 1$). The integral that defines the local Green's function [Eq. (6)] can now be performed analytically, to yield

$$G(z) = \frac{\tilde{z}}{2t^{*2}} - \text{sgn}[\text{Im}(\tilde{z})] \frac{\sqrt{\tilde{z}^2 - 4t^{*2}}}{2t^{*2}}, \quad (9)$$

with $\tilde{z} = z + \mu - \Sigma(z)$, and z an arbitrary complex number.

In addition, one can also calculate susceptibilities for charge-density-wave or spin-density-wave orders. When these susceptibilities become infinite, the system has a second-order phase transition to an ordered state. We will not describe in detail how to calculate such susceptibilities here (a discussion has appeared already[12,11]), because the system never underwent any second-order phase transitions for all of the parameters we considered in this work.

III. RESULTS

We begin our discussion in the low-conduction-band-density limit ($n_d \rightarrow 0$) which leads to the "excitonic phase" considered by Ramirez, Falicov, and Kimball.[1] The excitons (bound electron-hole pairs which do not contribute to the dc conductivity) can be shown to form when U is larger than a critical value U_c. The critical value of U was determined by solving the single-exciton problem (in the hole picture),[1] where the Hamiltonian in Eq. (1) with one hole and one electron produces a bound state just below the conduction band. This occurs when the Green's function, evaluated on the real axis at the lower band edge, is equal to $-1/U$, or

$$\frac{1}{U_c} = \int_{-2}^{2} d\epsilon \frac{D(\epsilon)}{\epsilon + 2 + i\delta} = -G(-2 + i\delta) = 1, \quad (10)$$

for the infinite-coordination-number Bethe lattice after using Eq. (9). This is the same critical value of U at which the interacting density of states splits into two bands as $n_d \rightarrow 0$.[14] However, it is not clear whether this simple criterion for exciton formation (based on the single-exciton problem) is sufficient to create an excitonic insulator when the conduction-electron density is small, but finite. Thus we analytically determine the conduction-band Green's functions at finite temperatures and finite, but small electron density. We find the solutions separate into two regimes: a weak-coupling regime with $U < 1$ and a strong-coupling regime with $U > 1$.

The functional form of the self-energy is expressed as a function of the local Green's function and the d-electron concentration as[12]

$$\Sigma(\omega) = \frac{U}{2} - \frac{1}{2G(\omega)}$$
$$\times \{1 \pm \sqrt{1 + 2(1 - 2n_d)UG(\omega) + U^2 G^2(\omega)}\}, \quad (11)$$

where the sign is chosen for each frequency in such a way that the self-energy is analytic. Since we are interested in the insulating phase ($n_d \rightarrow 0$), we expand Eq. (11) about $n_d = 0$ by factoring out $(1 + UG)^2$ from the square root. If $U < 1$

(single-band regime), the factor $1+UG(\omega)$ never vanishes [from Eq. (10)], so we must choose the negative sign, in order to have a vanishing self-energy as $U\rightarrow 0$. If $U>1$ (two-band regime), we must choose the positive sign for $\omega < -2$ in order to ensure analyticity (i.e., the right sign for Im Σ), since $1+UG(\omega)$ is negative there. Consequently, we choose the minus sign for $U<1$ (single-band regime) and the plus sign for $U>1$ (two-band regime). Hence, in the limit where $|n_d U/[1+UG(\omega)]|\ll 1$, we have

$$\Sigma(\omega;n_d\rightarrow 0)=\begin{cases} U-\dfrac{n_d U}{1+UG(\omega)} & \text{if } U<1 \\[3mm] -\dfrac{1}{G(\omega)}+\dfrac{n_d U}{1+UG(\omega)} & \text{if } U>1. \end{cases}$$
(12)

Combining this with Eq. (9), we obtain the functional form of the local Green's function for each case. The resulting form is

$$G(\omega;n_d\rightarrow 0)=\begin{cases} G_{\text{non}}(\omega-U)-\dfrac{n_d U G_{\text{non}}(\omega-U)}{[1+UG_{\text{non}}(\omega-U)]\sqrt{(\omega-U)^2-4}} & \text{if } U<1 \\[4mm] -\dfrac{1}{U}+n_d G_{\text{non}}\left(\omega+\dfrac{1}{U};t^*\rightarrow\sqrt{n_d}t^*\right) & \text{if } U>1, \end{cases}$$
(13)

where we used the noninteracting form of the Green's function, found in Eq. (9), with $\widetilde{z}\rightarrow\omega+i\delta$. Note that the Green's function for $U>1$ represents only the split-off lower band.

When $U<1$, as expected, the Green's function has just a perturbed form from the noninteracting Green's function which is shifted by U, $G_{\text{non}}(\omega-U)$, and the gap to particle-hole excitations at $T=0$ remains as Δ. Thus the number of the conduction electrons n_d at finite temperatures is activated as in a semiconductor with a fixed energy gap. On the other hand, when $U>1$, the Green's function for the lower band becomes a band-narrowed $(t^*\rightarrow\sqrt{n_d}t^*)$ noninteracting Green's function[14] (plus a constant shift by $-1/U$), and has a weight n_d (the upper band has weight $1-n_d$). The unit-charge condition $(n_f+n_d=1)$ implies that this lower band is always half-filled, and the energy of the system (in the limit $T\rightarrow 0$ and n_d fixed) becomes

$$E_{\text{insulator}}\rightarrow(1-n_d)E_f+2\int_{-(1/U)-2\sqrt{n_d}}^{-1/U}d\epsilon$$
$$\times\epsilon\left[-\frac{1}{\pi}\text{Im}G(\omega;n_d\rightarrow 0,U>1)\right]$$
$$=U-\Delta-2+n_d\left(\Delta+2-U-\frac{1}{U}-\frac{8}{3\pi}\sqrt{n_d}\right),$$
(14)

where the factor of 2 in the integral arises from the spin degeneracy of the conduction band. We will see below that this strong-coupling phase is not an excitonic insulator, as believed by Ramirez, Falicov, and Kimball[1] for any finite temperature. It is interesting to note, though, that this insulating limit [where all of the electrons lie in the localized states with an empty conduction band $(n_d=0)$] is noninteracting at $T=0$, because there are no conduction electrons, and hence the localized electrons do not feel any Coulomb repulsion. The ground-state energy is $E_{\text{insulator}}=U-\Delta-2$. The gap to particle-hole excitations (at $T=0$ for the insulating phase) is Δ for $U<1$, but decreases toward zero as $\Delta +2-U-(1/U)$ for $U\geq 1$. At the critical value of U, where the gap closes, one can see from the insulator energy in Eq.

(14) that the insulating phase is unstable if $\Delta+2-U-(1/U)<0$ (or equivalently if $U>1+(\Delta/2)+\sqrt{\Delta[1+(\Delta/4)]}$) since the ground-state energy is lowered for small, but nonzero n_d.

There is another phase of the FKM that is also noninteracting. It is the metallic phase (for large U) where the electrons (one per site) in the valence band are all promoted to the conduction band $(n_d=1)$. In this case, the conduction band "feels" no Coulomb repulsion, because there are no f electrons to scatter them. Thus the system is characterized by the half-filled noninteracting conduction band. The energy of this metallic phase is then

$$E_{\text{metal}}=2\int_{-2}^{0}d\epsilon\,\epsilon\,D(\epsilon)=-\frac{8}{3\pi}.$$
(15)

Therefore, at $T=0$, there is a transition from an insulating ground state to a metallic ground state when $E_{\text{metal}}<E_{\text{insulator}}$, or

$$U>\Delta+2-\frac{8}{3\pi}.$$
(16)

Surprisingly, there is a small region of U, $\Delta+2-(8/3\pi)>U>1+(\Delta/2)+\sqrt{\Delta[1+(\Delta/4)]}$, and a small region of the gap energy $0<\Delta<(3\pi/8)[1-(8/3\pi)]^2\approx 0.026\,923$, where the ground state is *neither metallic nor insulating*. In this nontrivial region, there must exist either an intermediate-valence state, or a charge-density-wave-ordered insulator. Detailed studies in this regime will appear in a future publication.

In this paper, we are interested in examining the discontinuous phase transitions between states that are connected either to the insulating phase or to the metallic phase as $T\rightarrow 0$. So we choose the bare gap to be large enough ($\Delta=1$), in order to be sufficiently far from any intermediate valence or charge-density-wave ordered phases. We vary U, and for each value of U calculate the thermodynamic properties of the system. We expect interesting behavior to occur for U close to the metal-insulator transition point at $T=0$, or
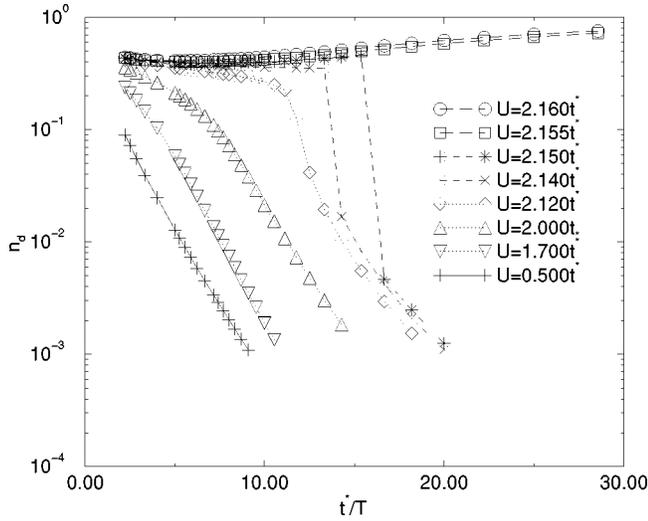
FIG. 1. Number density of the conduction electrons n_d (on a logarithmic scale) plotted as a function of t^*/T for various values of U, where $\Delta = 1.0t^*$. Four different regimes are shown: (a) the metallic regime for all T ($U = 2.160t^*$ and $2.155t^*$); (b) the discontinuous metal-insulator-transition regime ($U = 2.150t^*$ and $2.140t^*$); (c) the continuous metal-insulator-transition regime ($U = 2.120t^*$, $2.000t^*$, and $1.700t^*$); and (d) the semiconducting regime ($U = 0.500t^*$).

$U \approx 2.151\,17$ [from Eq. (16)]. We also expect simple semiconducting behavior (with a gap Δ) to occur for $U < 1$.

In Fig. 1, we present our numerical results of n_d (vertical axis) as a function of $1/T$ (horizontal axis) for different values of U, when $\Delta = 1$. The conduction-band filling n_d is plotted on a logarithmic scale, so that the linear behavior in $1/T$ indicates activated carriers as in a semiconductor (see $U = 0.5$). On the other hand, the metallic limit obtained in Eq. (16) for $U > 2.151\,17$ also agrees with the numerical results, since the system remains metallic for all temperatures. Figure 1 shows both discontinuous ($U = 2.150$ and 2.140) and continuous ($U = 2.120$, 2.000, and 1.700) metal-insulator transitions. In the discontinuous transitions, the conduction-electron concentration n_d follows the metallic solutions at high temperatures, but drops to the insulating solution at a sharp transition temperature. (For $U = 2.150$, n_d drops by two orders of magnitude as the temperature changes by $\sim 10^{-3}$.) This electron concentration can be viewed as an approximation of the electrical conductivity, if we assume that in a real material there is also static disorder (from defects, impurities, etc.) so that at low temperatures the relaxation time τ approaches a constant and the conductivity is proportional to $n_d \tau$. We also calculate the *intrinsic* conductivity of the FKM below, assuming that all of the scattering of the conduction electrons arises from the localized f electrons.

The conduction-electron density of states, $A(\omega) = -(1/\pi)\mathrm{Im}G(\omega)$ (where ω is measured from the chemical potential μ) provides additional information about the metal-insulator transitions. We calculate it by solving for the Green's functions on the real axis. Figure 2 plots $A(\omega)$ for some of the representative cases from Fig. 1 at various temperatures: (a) the metallic regime ($U = 2.160$), where the half-filled lower band increases in size as the temperature decreases; (b) the discontinuous metal-insulator-transition re-
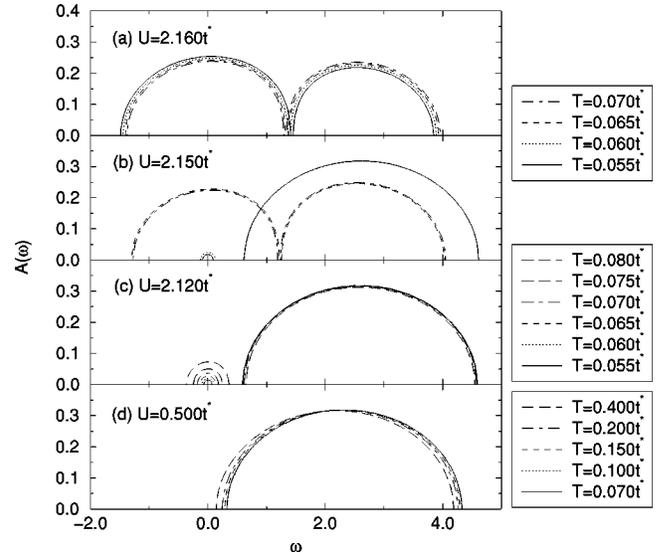
FIG. 2. Density of states $A(\omega)$ at different temperatures for some representative cases of Fig. 1: (a) $U = 2.160t^*$, (b) $U = 2.150t^*$, (c) $U = 2.120t^*$, and (d) $U = 0.500t^*$ when $\Delta = 1.0t^*$. Here the energy ω is measured from the chemical potential μ (i.e., the Fermi level lies at $\omega = 0$).

gime, where the corresponding density of states in the lower band discontinuously collapses to the insulating phase at a critical temperature T_c ($0.060 < T_c < 0.065$ for $U = 2.150$); (c) the continuous metal-insulator-transition regime, where the lower band is continuously reduced as the temperature decreases; and (d) the semiconducting regime, where n_d has an activated behavior, and $A(\omega)$ displays a finite pseudogap from the perturbed single conduction band. Note that all of the temperature dependence of the interacting density of states arises from the temperature dependence of n_d, since the FKM with fixed n_d and n_f has a *temperature-independent* density of states.[14]

The exact solution of the FKM also allows the optical conductivity to be calculated from the following formula:[15,16]

$$\sigma(\omega) = \sigma_0 \pi \int_{-\infty}^{\infty} d\omega' \int_{-\infty}^{\infty} d\epsilon (4 - \epsilon^2) D(\epsilon) A(\epsilon, \omega')$$

$$\times A(\epsilon, \omega' + \omega) \frac{f(\omega') - f(\omega' + \omega)}{\omega}, \qquad (17)$$

where $f(\omega)$ is the Fermi distribution function, and σ_0 gives the conductivity unit. This formula for the conductivity is the result of a derivation,[16] which corrects the error of using the formula derived for the hypercubic lattice[15] by adding the factor $4 - \epsilon^2$. (In most cases, the addition of this factor does not significantly change the qualitative features of the conductivity.) The dc conductivity σ_{dc} is then found from Eq. (17) in the limit $\omega \to 0$.

Substituting the spectral function into Eq. (17), and assuming that the self-energy has a negligible frequency dependence near the Fermi level, produces the following limiting form for the dc conductivity with $T = 0$ and $n_d \to 0$ (i.e., for the insulating phase):
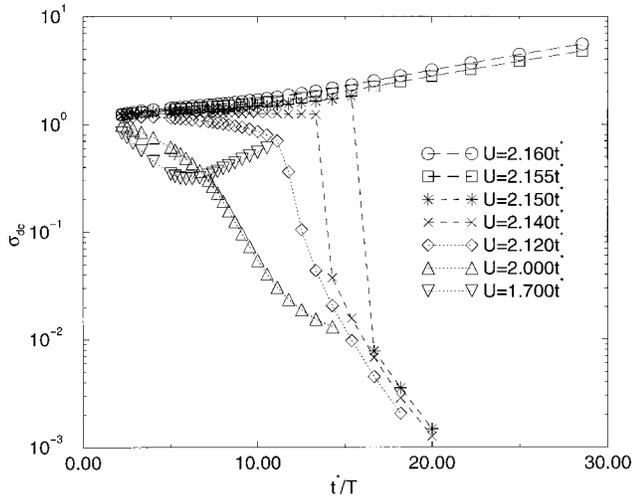
FIG. 3. The *intrinsic* dc conductivity σ_{dc} of the Falicov-Kimball model as a function of t^*/T for the corresponding values of U shown in Fig. 1. Note how the intrinsic conductivity diverges for low temperature and low electron concentration, as described in the text.

$$\sigma_{dc}(T=0; n_d \to 0) = -\frac{\sigma_0 \{4 - [\mu - \mathrm{Re}\Sigma(0)]^2\}^{3/2}}{4\pi\mathrm{Im}\Sigma(0)}, \tag{18}$$

which is proportional to the product of the *intrinsic* relaxation time ($\sim 1/\mathrm{Im}\Sigma$) and the "conductivity" density of states at the Fermi surface. Calculating the self-energy $\Sigma(0)$ from Eqs. (12) and (13) with the proper chemical potential for each case ($\mu = U - 2 + \delta\mu$ for $U < 1$, and $\mu = -1/U$ for $U > 1$), we obtain

$$\sigma_{dc}(T=0; n_d \to 0) = \begin{cases} \sigma_0 \left(\frac{1}{U} - 1\right)^2 \left(\frac{9}{2\pi n_d}\right)^{1/3} & \text{if } U < 1 \\[2ex] \dfrac{\sigma_0 \sqrt{4 - \left(1 + \dfrac{1}{U}\right)^{3/2}}}{4\pi(U-1)\sqrt{n_d}} & \text{if } U > 1. \end{cases} \tag{19}$$

Thus the *intrinsic* σ_{dc} (which is obtained from the pure electronic system of the model) in the insulating phase at $T = 0$ *actually diverges* as $n_d \to 0$. Even in the strong-coupling regime ($U > 1$), the electron-hole excitations do not bind to form an excitonic insulator. [The dc conductivity actually diverges faster ($\sim 1/\sqrt{n_d}$) than in the weak-coupling regime ($\sim 1/n_d^{1/3}$) as $n_d \to 0$.] Therefore, the excitonic phase considered by Ramirez, Falicov, and Kimball[1] does not exist on the infinite-coordination-number Bethe lattice. Moreover, the divergence of the *intrinsic* σ_{dc} occurs because the relaxation time τ increases more rapidly than the "conductivity" density of states at the Fermi surface decreases when $n_d \to 0$ [see Eq. (18)]. But in a real material the relaxation time can never diverge, because there always exists some static disorder which forces the relaxation time to approach a constant at low temperature. Hence the dc conductivity will approach zero as $n_d \to 0$ in any real material. In Fig. 3, we present our numerical results of the *intrinsic* dc conductivity as a func-
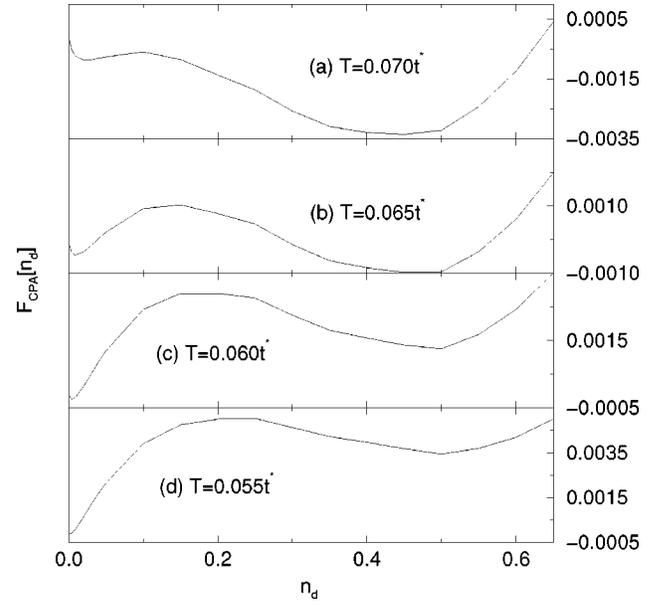
FIG. 4. Free energy as a function of the number density of the conduction electrons near the first-order transition temperature. The parameters are $\Delta = 1.0t^*$ and $U = 2.150t^*$. The free energy has two local minima with the global minimum switching between these local minima as temperature varies through the critical temperature ($0.060 < T_c < 0.065$), indicating a first-order transition.

tion of $1/T$ for the corresponding values of U in Fig. 1. The conductivity in Fig. 3, with moderate n_d, appears to be proportional to the electron concentration shown in Fig. 1, but σ_{dc} starts to diverge as temperature decreases, and n_d becomes small enough for the limiting form in Eq. (19) to hold. (For example, see the cases $U = 2.000$ and 1.700. For $U = 0.500$, σ_{dc} is always in the low-density limit.) Thus the relevant dc conductivity for a real material is approximated better by $n_d \tau$ with a constant relaxation rate τ, rather than using the *intrinsic* dc conductivity.

Finally, we examine the discontinuous metal-insulator transitions in more detail to show that they are indeed first-order phase transitions. We do this by employing the CPA formalism to calculate the trial free energy as a function of n_d (or, equivalently, $n_f = 1 - n_d$) at different temperatures. The minimum of the trial free energy determines the thermodynamic electron density. We first solve the real-axis self-consistent equations for each value of $0 < n_d < 1$ to find the interacting density of states $A(\omega)$. Then we evaluate the exact form of the free energy in Eq. (7) as a function of n_d, and repeat the process for different temperatures. We present our results near the critical temperature in Fig. 4, where $U = 2.150$ and $\Delta = 1$. The free energy has a double minimum near the critical temperature ($0.060 < T_c < 0.065$), and, as the temperature is lowered, the conduction-electron density discontinuously changes as the global minimum switches between the two local minima (indicating a first-order transition). At T_c, where the two minima are degenerate, the system exhibits phase coexistence between the insulating and metallic phases. In the region where the change in n_d is continuous, the free energy does not have multiple minima, but rather the minimum of the free energy varies smoothly with n_d as the temperature is changed. We also verified that

the minimum of the CPA form of the free energy agrees with the form for the free energy determined by Brandt and Mielsch.[12]

IV. CONCLUSIONS

In conclusion, we have exactly solved the spin-$\frac{1}{2}$ Falicov-Kimball model on an infinite-coordination-number Bethe lattice, which is shown to have both continuous and discontinuous (first-order) charge-transfer metal-insulator transitions. By being able to solve the model exactly, we have clarified the theoretical controversy of the model, and have proven that the model does display first-order metal-insulator transitions. The simplicity of the model, based on only the electronic system (which has both a localized and a conduction band), emphasizes the fact that the electronic system itself (the Coulomb interaction between a conduction electron and a localized electron) can cause dramatic discontinuous charge-transfer metal-insulator transitions without requiring other effects (such as phonons). We expect our results to continue to hold in three dimensions, and to have applications to real materials such as NiI_2.

We also found that the intrinsic conductivity (determined by the scattering of the d electrons off the f electrons) actually diverges for the ''insulating'' phases with $n_d \to 0$, because the relaxation time grows faster than the density of states at the Fermi level decreases. In a real material, the conductivity will go to zero as $n_d \to 0$ though, because the relaxation time is bounded by scattering off of impurities. We also discovered a small region of parameter space that possesses intermediate-valence or charge-density-wave order. Further studies of this region are currently underway.

ACKNOWLEDGMENTS

[1] L. M. Falicov and J. C. Kimball, Phys. Rev. Lett. **22**, 997 (1969); R. Ramirez, L. M. Falicov, and J. C. Kimball, Phys. Rev. B **2**, 3383 (1970).

[2] N. F. Mott, Proc. Phys. Soc. London, Sect. A **62**, 416 (1949); Prog. Metal Phys. **3**, 76 (1956); Can. J. Phys. **34**, 1356 (1956); Philos. Mag. **6**, 281 (1961); J. Hubbard, Proc. R. Soc. London, Ser. A **281**, 401 (1964); B. H. Brandow, Adv. Phys. **26**, 651 (1977).

[3] P. W. Anderson, Phys. Rev. **109**, 1492 (1958); E. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan, Phys. Rev. Lett. **42**, 673 (1979); P. A. Lee and T. V. Ramakrishnan, Rev. Mod. Phys. **57**, 287 (1985).

[4] M. P. Pasternak, R. D. Taylor, A. Chen, C. Meade, L. M. Falicov, A. Giesekus, R. Jeanloz, and P. Y. Yu, Phys. Rev. Lett. **65**, 790 (1990); A. L. Chen, P. Y. Yu, and R. D. Taylor, *ibid.* **71**, 4011 (1993).

[5] J. K. Freericks and L. M. Falicov, Phys. Rev. B **45**, 1896 (1992).

[6] C. E. T. Goncalves da Silva and L. M. Falicov, J. Phys. C **5**, 906 (1972).

[7] M. Plischke, Phys. Rev. Lett. **28**, 361 (1972); D. K. Ghosh, Solid State Commun. **18**, 1377 (1977).

[8] D. I. Khomskii, in *Quantum Theory of Solids*, edited by I. M. Lifshits (Mir, Moscow, 1982); P. Farkasovsky, Phys. Rev. B **52**, R5463 (1995).

[9] J. K. Freericks, Phys. Rev. B **47**, 9263 (1993).

[10] T. Kennedy and E. H. Lieb, Physica A **138**, 320 (1986); E. H. Lieb, *ibid.* **140**, 240 (1986); P. G. J. van Dongen and D. Vollhardt, Phys. Rev. Lett. **65**, 1663 (1990).

[11] J. K. Freericks and V. Zlatić (unpublished).

[12] U. Brandt and C. Mielsch, Z. Phys. B **75**, 365 (1989); **79**, 295 (1990); U. Brandt, A. Fledderjohann, and G. Hülsenbeck, *ibid.* **81**, 409 (1990); U. Brandt and C. Mielsch, *ibid.* **82**, 37 (1991).

[13] M. Jarrell, Phys. Rev. Lett. **69**, 168 (1992).

[14] P. G. J. van Dongen and C. Leinung, Ann. Phys. (Leipzig) **6**, 45 (1997).

[15] Th. Pruschke, D. L. Cox, and M. Jarrell, Phys. Rev. B **47**, 3553 (1993).

[16] Woonki Chung and J. K. Freericks (unpublished).