Critical behavior of the density of states at the metal-insulator transition

D. Belitz
Department of Physics and Materials Science Institute, University of Oregon, Eugene, Oregon 97403

T. R. Kirkpatrick
Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742
(Received 12 May 1993)

We perform a thorough analysis of the critical behavior of the density of states at the metal-insulator transition within the framework of the field-theoretic renormalization group. Of the previous, conflicting results obtained by resumming the perturbation theory, and by frequency-momentum-shell renormalization-group methods, we agree with the one yielding conventional power-law behavior. Emphasis is placed on a detailed technical derivation of this result, and on peculiarities in the renormalization-group description that arises from the long-ranged nature of the Coulomb interaction, and from the presence of multiple time scales in the problem.

I. INTRODUCTION

Interacting electrons in a random environment of static scatterers display a zero-temperature phase transition from an electric conductor to an insulator if the effective disorder is increased beyond a critical value.[1] In contrast to the corresponding Anderson transition of noninteracting electrons, where the single-particle or tunneling density of states (DOS) is uncritical,[2] this Anderson-Mott transition is characterized by a critical DOS at the Fermi level. The vanishing of the DOS at the metal-insulator transition (MIT) has a precursor in the so-called Coulomb anomaly of the DOS in the metallic phase,[3] and presumably finds its continuation in the Coulomb gap in the insulating phase.[4] The criticality of the DOS has been observed in a variety of materials.[5]

Our current theoretical understanding of the MIT is as follows. There are four distinct universality classes.[6,1] They are represented by systems with magnetic impurities (MI), systems in a magnetic field (MF), systems with spin-orbit scattering (SO), and the generic case of no spin-dependent external fields (G). Physically, the first two classes (MI and MF) are distinct from the last two because magnetic impurities or magnetic fields suppress the particle-particle or Cooper scattering and interaction channels, while the Cooper channel is present and influences the critical behavior in the SO and G universality classes. Technically, a generalization of Wegner's field-theoretic approach to the Anderson transition[7] has proved very successful in describing the MIT for the classes MI and MF.[8] The asymptotic critical behavior at the transition is characterized by three independent exponents, ν, z, and β. ν describes the divergence of the correlation length ξ according to $\xi \sim t^{-\nu}$. z describes the vanishing of the (dominant) frequency (Ω) or temperature (T) scale according to $\Omega \sim T \sim \xi^{-z}$. Finally, the DOS at the Fermi level vanishes as $N(\varepsilon_F) \sim t^{\beta}$. Here t denotes the dimensional distance from the critical point at zero temperature and frequency. The exponents ν, z, and β were calculated in Ref. 8 in lowest order in an ε expansion about $D=2$. All other critical exponents are related to ν, z, and β by scaling laws. The same problem was treated in Ref. 9 by means of a different method. The results agree with those of Ref. 8 except for the DOS, for which Ref. 9 found not a power law, but rather $\ln N(\varepsilon_F) \sim -(\ln t)^2$, i.e., $\beta = \infty$. The reason for this discrepancy is not easy to see, and the point has never been clarified. Technically, the calculation for the DOS in Ref. 8 is on less firm ground than the remainder of the paper. The author at this point abandons the frequency-momentum-shell renormalization-group (RG) methods which he employed for all other quantities and resorts to a direct exponentiation of perturbation theory. On the other hand, the validity of the frequency-momentum-shell RG employed in Ref. 9 is not obvious either, mainly since the limits of criticality and $D \to 2$ do not commute in the case of the DOS (see the discussion in Sec. IV C 1).

It is the purpose of the present paper to clarify this point. Using field-theoretic methods we will derive the scaling properties of the DOS and show that Finkel'stein's result, Ref. 8, is the correct one. The method used will also shed light on some subtle points related to the long-ranged nature of the Coulomb interaction and how to deal with it in the framework of the RG. We also discuss the consequences of multiple time scales in the MIT problem.

We will purposefully restrict ourselves to the MI and MF universality classes. For the SO and G classes, where cooperons are present, conflicting results have been obtained[10,11] for quantities other than the DOS, and no accepted solution of the problem exists. However, some technical aspects of the DOS problem are related to and relevant for what we believe will be the solution of the cooperon problem.[12] These points will be mentioned in the discussion.

The remainder of this paper is organized as follows. In Sec. II we first recall the basic structure of the theory and the loop expansion. The relevant one-point vertex function is then given in perturbation theory to one-loop order. In Sec. III the theory is renormalized and a Callan-

Symanzik equation for the DOS is derived. The solution of this equation in Sec. IV yields the leading scaling behavior as well as corrections to scaling. The paper is concluded with a discussion of our techniques and results, and of the light they shed on some unsolved aspects of the MIT problem.

II. STRUCTURE OF THE THEORY, AND THE LOOP EXPANSION

A. The generalized nonlinear σ model

Our starting point is the fermionic generalization[6,8] of Wegner's matrix nonlinear σ model[7] to the case of interacting electrons. The action reads

$$S[Q]=\frac{-1}{2G}\int d\mathbf{x}\,\mathrm{tr}\left[\nabla Q(\mathbf{x})+\frac{e}{c}\mathbf{A}(\mathbf{x})[Q(\mathbf{x}),\tau_3\otimes s_0]\right]^2 +2H\int d\mathbf{x}\,\mathrm{tr}[\Omega Q(\mathbf{x})]+b\int d\mathbf{x}\,\mathrm{tr}[\tau_3\otimes s_3 Q(\mathbf{x})] -\frac{\pi N_F}{6\tau_s}\int d\mathbf{x}\sum_{i=1}^{3}\mathrm{tr}[\Sigma_i Q(\mathbf{x})]^2+S_{\mathrm{int}}[Q]\ . \tag{2.1}$$

The matrix field $Q(\mathbf{x})$ comprises products of fermionic degrees of freedom, and depends on replica labels α,β ($\alpha,\beta=1,2,\ldots,N$) and on Matsubara frequency labels n,m. The number of replicas N will be set to zero after calculations, since the replica trick is employed in dealing with the quenched disorder. Q is traceless, Hermitian, and unitary, and can be parametrized,

$$Q=\begin{array}{c|c} \begin{matrix} m\geq 0 & \qquad m<0 \end{matrix} & \\ \hline \begin{pmatrix} (1-qq^+)^{1/2}-1 & q \\ q^+ & -(1-q^+q)^{1/2}+1 \end{pmatrix} & \begin{matrix} n\geq 0 \\ n<0\ . \end{matrix} \end{array} \tag{2.2a}$$

Here the q are matrices with spin-quaternion valued elements $q_{nm}^{\alpha\beta}$; $n=0,1,\ldots$; $m=-1,-2,\ldots$. They are conveniently expanded in a spin-quaternion basis,

$$q_{nm}^{\alpha\beta}(\mathbf{x})=\sum_{r=0}^{3}\sum_{i=0}^{3}{}_i^r q_{nm}^{\alpha\beta}(\mathbf{x})\tau_r\otimes s_i\ , \tag{2.2b}$$

where $\tau_0=s_0=1_2$ with 1_2 the unit 2×2 matrix, and $\tau_j=s_j=-i\sigma_j$ ($j=1,2,3$) with σ_j the Pauli matrices. The matrices τ_r and s_i describe the particle-hole and spin degrees of freedom, respectively. The coupling constant G, which plays a role analogous to the temperature in a finite-temperature phase transition, is $G=8e^2/\pi\sigma$ with σ the bare (i.e., self-consistent Born) conductivity. e is the electron charge, and we use units such that $\hbar=1$. G is a measure of the disorder strength. $H=\pi N_F/4$ with N_F the bare DOS at the Fermi level is a frequency coupling parameter, and $\Omega_{nm}^{\alpha\beta}=\delta_{\alpha\beta}\delta_{nm}\omega_n\tau_0\otimes s_0$, with $\omega_n=2\pi T(n+\frac{1}{2})$ a fermionic Matsubara frequency, is the external frequency matrix. $\mathbf{A}$ denotes the vector potential, and $b=\pi N_F g_L\mu_B B/2$ is proportional to the corresponding magnetic field B. g_L is the Landé factor, and μ_B the Bohr magneton. τ_s is the spin-flip time due to magnetic impurities, and $(\Sigma_i)_{nm}^{\alpha\beta}=\delta_{\alpha\beta}\delta_{nm}\tau_3\otimes s_i$.

In the absence of S_{int} this model describes the localization of noninteracting electrons.[7,6] The motivation for and derivation of the model have been discussed many times.[6,8,1] Let us mention only that the critical behavior at the MIT, as at any continuous phase transition, is governed by the slow modes of the system. The slow modes described by Eq. (2.1) are the diffusive ones that are related to the conservation of particle number, spin, and energy density. The basic underlying assumptions are that (a) there are no other slow modes, and (b) all massive excitations are irrelevant for the critical behavior. An inspection of Eq. (2.1) shows that in the presence of either a magnetic field ($\mathbf{A}$, $b\neq 0$) or magnetic impurities ($1/\tau_s\neq 0$) the modes with $r=1,2$ (which describe the particle-particle or cooper channel) are massive. According to the above-stated assumptions they can therefore be neglected. In addition, magnetic impurities introduce a mass into all spin-triplet models ($i=1,2,3$), while with only a magnetic field present only two of the three spin-triplet modes ($i=1,2$) become massive.

The term S_{int} in Eq. (2.1) describes the electron-electron interaction. It is a four-fermion term and therefore quadratic in Q. As mentioned in the Introduction, we are restricting ourselves to systems with either $b\neq 0$ (universality class MF) or $1/\tau_s\neq 0$ (universality class MI). According to the above discussion, we therefore have to consider only the interaction in the particle-hole channel ($r=0,3$). The corresponding part of S_{int} can be written[8,1]

$$S_{\mathrm{int}}[Q]=\frac{-\pi T}{4}\int d\mathbf{x}\,d\mathbf{y}\sum_{\substack{n_1n_2\\n_3n_4}}\delta_{n_1+n_3,n_2+n_4}\sum_{\alpha}\sum_{r=0,3}(-)^r\left[\mathrm{tr}[\tau_r\otimes s_0 Q_{n_1n_2}^{\alpha\alpha}(\mathbf{x})]K_s(\mathbf{x}-\mathbf{y})\,\mathrm{tr}[\tau_r\otimes s_0 Q_{n_3n_4}^{\alpha\alpha}(\mathbf{y})] +\sum_{i=1}^{3}\mathrm{tr}[\tau_r\otimes s_i Q_{n_1n_2}^{\alpha\alpha}(\mathbf{x})]K_t\delta(\mathbf{x}-\mathbf{y})\,\mathrm{tr}[\tau_r\otimes s_i Q_{n_3n_4}^{\alpha\alpha}(\mathbf{y})]\right]\ . \tag{2.3a}$$

Here T is the temperature, and K_t is the spin-triplet interaction amplitude which can be considered pointlike in our long-wavelength theory. K_s, the spin-triplet interaction amplitude, is the statically screened Coulomb interaction. Its Fourier transform is

$$K_s(\mathbf{p})=\frac{-\widetilde{H}}{p^{D-1}+\kappa_D^{D-1}}[p^{D-1}F_0^s+\kappa_D^{D-1}(1+F_0^s)]\ . \tag{2.3b}$$

Here F_0^s is a Landau parameter, $\tilde{H}=H/(1+F_0^s)=(\pi/8)\partial n/\partial\mu$ is proportional to the thermodynamic density susceptibility $\partial n/\partial\mu$, and $\kappa_D=(\pi 2^{D-1}e^2\partial n/\partial\mu)^{1/(D-1)}$ is the screening wave number in D dimensions. Notice that $K_s(p=0)+H=0$. We therefore must keep the p dependence of K_s lest we lose the generic structure of the density propagator, cf. Eq. (2.4b) below.

B. Propagators, vertex functions, and the loop expansion

The next step is to expand the action in powers of q by means of Eq. (2.2a). At the Gaussian or zero-loop level (i.e., keeping only the terms quadratic in q) one can express arbitrary q-correlation functions in terms of three propagators,[13,1]

$$\mathcal{D}_n(p)=[p^2+GH\Omega_n]^{-1}\ , \tag{2.4a}$$

$$\mathcal{D}_n^s(p)=\{p^2+G[H+K_s(p)]\Omega_n\}^{-1}\ , \tag{2.4b}$$

$$\mathcal{D}_n^t(p)=[p^2+G(H+K_t)\Omega_n]^{-1}\ . \tag{2.4c}$$

Here $\Omega_n=2\pi Tn$ $(n=0,1,\ldots)$ is a bosonic Matsubara frequency. Terms of higher order in q in the action can be taken into account perturbatively by means of a systematic loop expansion.[14] In our case this amounts to an expansion in powers of G.

The quantity of central interest for this paper is the single-particle DOS at zero temperature as a function of disorder and energy or temperature, which we will denote by $N(\varepsilon)=N(i\Omega_n\to|\varepsilon-\varepsilon_F|+i0)$. At zero-loop order, $N(\varepsilon)\equiv N_F$. In general, N can be expressed in terms of a one-point correlation function,[8]

$$N(\Omega_n)=N_F\langle {}_0^0Q_{nn}^{\alpha\alpha}(\mathbf{x})\rangle\ , \tag{2.5}$$

where the average is to be taken with the full action $S[Q]$. The corresponding vertex part is the one-point vertex function, which we denote by $\Gamma^{(1)}$. To one-loop order one easily obtains for $\Gamma^{(1)}$ (Ref. 1)

$$\Gamma^{(1)}(\Omega_n;G,H,\kappa_D)=1+\frac{G}{8}\sum_{l=1}^{\infty}\sum_{\mathbf{p}}\frac{1}{l+n}\Delta\mathcal{D}_{l+n}^s(p)+O(G^2)\ , \tag{2.6a}$$

where $\Delta\mathcal{D}_l^s(p)=\mathcal{D}_l^s(p)-\mathcal{D}_l(p)$. In what follows we will prefer a cutoff regularized theory over the dimensional regularization employed elsewhere.[13,1] Accordingly, we introduce a momentum cutoff Λ and a frequency cutoff $\Omega_0=\Lambda^2/GH$. We also analytically continue to real frequencies and find

$$\Gamma^{(1)}(\Omega;G,H,\kappa_D;\Lambda)=1+\frac{\bar{G}}{8}GH\kappa_D^{1+\varepsilon}\int_{\Omega}^{\Omega_0}d\omega\int_0^{\Lambda}dp\frac{1}{p^2+GH\omega}\,\frac{1}{\kappa_D^{1+\varepsilon}p^{1-\varepsilon}+G\tilde{H}\omega}+O(G^2)\ , \tag{2.6b}$$

where $\bar{G}=GS_D/(2\pi)^D$ with S_D the surface of the D-dimensional unit sphere. Equation (2.6b) holds for the MI class. In the case of the MF class, there is an additional term with $\Delta\mathcal{D}^s$ in Eq. (2.6a) replaced by $\Delta\mathcal{D}^t=\mathcal{D}^t-\mathcal{D}$. Inspection shows that for $D\to 2$ the singlet contribution is more singular due to the well-known $\log^2$ singularity that exists in $D=2$ and was first discussed by Altshuler, Aronov, and Lee.[3] The one-loop spin-triplet contribution to the exponent β, to be determined in Sec. IV, is of the same order as the two-loop spin-singlet contribution. We therefore neglect it here. Equation (2.6b) gives the leading contribution to $\Gamma^{(1)}$ for both the MI and the MF universality classes.[15] Performing the frequency integration in Eq. (2.6b) yields

$$\Gamma^{(1)}(\Omega;G,H,\kappa_D;\Lambda)=1+\frac{\bar{G}}{8}\Lambda^{\varepsilon}\int_0^1\frac{dx}{x^{1-\varepsilon}-x^2(\tilde{H}/H)(\Lambda/\kappa_D)^{1+\varepsilon}}\ln\left[\frac{x^2+1}{x^2+\Omega GH/\Lambda^2}\,\frac{x^{1-\varepsilon}+(\Omega G\tilde{H}/\Lambda^2)(\Lambda/\kappa_D)^{1+\varepsilon}}{x^{1-\varepsilon}+(\tilde{H}/H)(\Lambda/\kappa_D)^{1+\varepsilon}}\right]\ . \tag{2.7}$$

Notice that for $\kappa_D=\infty$ and for $\Lambda\to\infty$ one has $\Gamma^{(1)}=1+(\bar{G}/4\varepsilon^2)[1+O(\varepsilon^2)]\Lambda^{\varepsilon}$ (the $1/\varepsilon^2$ prefactor reflects the strong singularity mentioned above) while for $\kappa_D<\infty$ the integral is finite even for $\Lambda\to\infty$. From a purely perturbative point of view, the cutoff is therefore not necessary. However, in the next sections we will obtain very useful information from studying the dependence of $\Gamma^{(1)}$ on Λ.

III. RENORMALIZATION

A. The renormalized one-point vertex function

We now apply the canonical field-theoretic renormalization scheme[14] to the one-point vertex function, Eq. (2.7). In doing so we *assume* that the theory is renormalizable with four renormalization constants: One each for the three coupling constants, G, H, and K_t, and one field or wave-function renormalization.[16] We define renormalized coupling constants g,h,

$$\bar{G}=\mu^{-\varepsilon}Z_g g\ , \tag{3.1a}$$

$$H=Z_h h\ . \tag{3.1b}$$

Here μ is the arbitrary RG momentum scale, and the renormalization constants Z_g and Z_h are functions of the renormalized coupling constants, and of Λ/μ and Λ/κ_D. The loop expansion corresponds to an expansion of Z_g and Z_h in powers of g,

$$Z_{g,h}=1+O(g) \ . \tag{3.1c}$$

If we define a renormalized screening wave number κ_D^R by

$$\kappa_D=\mu Z_\kappa \kappa_D^R \ , \tag{3.1d}$$

then the argument Λ/κ_D of $Z_{g,h}$ gets replaced by κ_D^R. The renormalized one-point vertex function is

$$\Gamma_R^{(1)}(\Omega;g,h,\kappa_D^R;\mu,\Lambda) = Z^{1/2}(g,h,\kappa_D^R,\Lambda/\mu)\Gamma^{(1)}(\Omega;G,H,\kappa_D;\Lambda) \ . \tag{3.2}$$

Notice that the field renormalization constant Z is dimensionless and can depend on Λ only through the combination Λ/μ. We eliminate the arbitrariness in the choice of Z by requiring the following normalization condition:

$$\Gamma_R^{(1)}(\Omega=\mu^D/gh;g,h,\kappa_D^R;\mu,\Lambda)=1 \ . \tag{3.3}$$

From Eqs. (3.2) and (3.3) we obtain for the field renormalization constant

$$Z(g,h,\kappa_D^R,\Lambda/\mu)=1-\frac{g}{4}\left[\frac{\Lambda}{\mu}\right]^{\varepsilon}\int_0^1 \frac{dx}{x^{1-\varepsilon}-x^2(\tilde{H}/h)(\Lambda/\mu)^{1+\varepsilon}/(\kappa_D^R)^{1+\varepsilon}} \times\ln\left[\frac{x^2+1}{x^2+(\mu/\Lambda)^2}\,\frac{x^{1-\varepsilon}+(\tilde{H}/h)(\mu/\Lambda)^{1-\varepsilon}/(\kappa_D^R)^{1+\varepsilon}}{x^{1-\varepsilon}+(\tilde{H}/h)(\Lambda/\mu)^{1+\varepsilon}/(\kappa_D^R)^{1+\varepsilon}}\right] , \tag{3.4}$$

and for the renormalized vertex function,

$$\Gamma_R^{(1)}(\Omega;g,h,\kappa_D^R;\mu,\Lambda)=1-\frac{g}{8}\int_0^{\Lambda/\mu}\frac{dx}{x^{1-\varepsilon}-x^2(\tilde{H}/h)/(\kappa_D^R)^{1+\varepsilon}}\ln\left[\frac{x^2+gh\Omega/\mu^D}{x^2+1}\,\frac{x^{1-\varepsilon}+(\tilde{H}/h)/(\kappa_D^R)^{1+\varepsilon}}{x^{1-\varepsilon}+(\tilde{H}/h)(gh\Omega/\mu^D)/(\kappa_D^R)^{1+\varepsilon}}\right] . \tag{3.5}$$

We see that $\Gamma_R^{(1)}$ is indeed finite as $\Lambda\to\infty$ with fixed renormalized coupling constants. For $\Lambda\gg\mu$ we have

$$\frac{d}{d\ln(\Lambda/\mu)}\Gamma_R^{(1)}=\frac{g}{8}\left[\frac{\mu}{\Lambda}\right]^{2-\varepsilon}\left[1-\frac{gh\Omega}{\mu^D}\right] \times[1+O(\mu/\Lambda)^{1+\varepsilon}] \ . \tag{3.6}$$

B. Callan-Symanzik equation

We now use standard arguments to derive a Callan-Symanzik (CS) equation for the bare one-point vertex function.[17] Note that the bare theory is what we are interested in, while the renormalized theory is an artifact that has no direct physical meaning (this is the opposite of the situation in high-energy physics). It is therefore natural to derive an equation directly for the bare vertex function.

Let us define dimensionless bare coupling constants g_0, h_0, and k_0 by

$$g_0=\bar{G}\Lambda^{\varepsilon} \ , \tag{3.7a}$$

$$h_0=H/\tilde{H} \ , \tag{3.7b}$$

$$k_0=\Lambda/\kappa_D \ . \tag{3.7c}$$

We note that the thermodynamic density susceptibility $\tilde{H}\sim\partial n/\partial\mu$ does not get renormalized.[8] The scaling properties of h_0 are therefore the same as those of H. κ_D, on the other hand, is given by $e^2\partial n/\partial\mu$, and the charge might get renormalized at a MIT. In general we therefore expect nontrivial scaling behavior of k_0. We will come back to this point.

By writing the total differential on the left-hand side of Eq. (3.6) explicitly, and using Eq. (3.2), we obtain a partial differential equation (PDE) for $\Gamma^{(1)}$,

$$\left\{\Lambda\frac{\partial}{\partial\Lambda}+\beta(x_0)\frac{\partial}{\partial g_0}+\zeta(x_0)\frac{\partial}{\partial h_0}+\rho(x_0)\frac{\partial}{\partial k_0}+\frac{1}{2}\gamma(x_0)\right\} \times\Gamma^{(1)}(\Omega;x_0;\Lambda)=O((\mu/\Lambda)^{2-\varepsilon}) \ . \tag{3.8}$$

Here $x_0=\{g_0,h_0,\ldots\}$ collectively denotes the bare coupling constants, and the RG functions β, ζ, etc. are defined as

$$\beta(x_0)=\Lambda\frac{dg_0}{d\Lambda} \ , \tag{3.9a}$$

$$\zeta(x_0)=\Lambda\frac{dh_0}{d\Lambda} \ , \tag{3.9b}$$

$$\rho(x_0)=\Lambda\frac{dk_0}{d\Lambda} \ , \tag{3.9c}$$

$$\gamma(x_0)=\Lambda\frac{d\ln Z}{d\Lambda} \ , \tag{3.9d}$$

with all derivatives taken at *fixed renormalized theory*. Strictly speaking, the RG functions will in general depend on μ/Λ as well as on x_0. However, since the bare theory cannot depend on μ, these dependencies must cancel against the (μ/Λ) dependence of the inhomogeneity on the right-hand side of Eq. (3.8). We therefore consistently neglect these (μ/Λ) dependencies. Thus we obtain the CS equation, i.e., Eq. (3.8) with the right-hand side replaced by zero, and the RG functions calculated in the limit $\Lambda/\mu\to\infty$.

C. The RG functions

Before we can solve the CS equation (3.8) we need the RG functions, Eqs. (3.9). Let us first consider $\gamma(x_0)$. From Eqs. (3.9d) and (3.4) we obtain, in the limit $\Lambda/\mu\to\infty$,

$$\gamma(x_0)=-\frac{g_0}{2\varepsilon}g\left[\frac{k_0^{1+\varepsilon}}{h_0}\right]+O(g_0^2)\ , \tag{3.10a}$$

with

$$g(z)=\varepsilon\int_0^1\frac{dx}{x^{1-\varepsilon}+z}\frac{1}{x^2+1}\ . \tag{3.10b}$$

For later reference we note that $g(z=0)=1+O(\varepsilon)$, and that for $D=2$ we have $\gamma(x_0)\sim g_0\ln(k_0/h_0)$.

For the functions $\beta(x_0)$ and $\zeta(x_0)$ we use the known results from the literature.[8,1] [For possible k_0 dependencies of these functions, see the discussion after Eq. (3.10′).] For the MI class one has

$$\beta_{\rm MI}(x_0)=\varepsilon g_0-g_0^2/4+O(g_0^3)\ , \tag{3.11a}$$

$$\zeta_{\rm MI}(x_0)=g_0h_0/8+O(g_0^2)\ . \tag{3.11b}$$

For the MF class one has

$$\beta_{\rm MF}(x_0)=\varepsilon g_0-\tfrac{1}{2}g_0^2\left[1-\frac{1}{2}\left[1+\frac{1}{\gamma_t^0}\right]l_t^0\right]+O(g_0^3)\ , \tag{3.12a}$$

with $\gamma_t^0=K_t/H$, and $l_t^0=\ln(1+\gamma_t^0)$, and

$$\zeta_{\rm MF}(x_0)=\tfrac{1}{8}g_0h_0(1-\gamma_t^0)+O(g_0^2)\ . \tag{3.12b}$$

γ_t^0 obeys a flow equation of its own,

$$\Lambda\frac{d\gamma_t^0}{d\Lambda}=-\tfrac{1}{8}g_0[1-(\gamma_t^0)^2]+O(g_0^2)\ . \tag{3.12c}$$

Note that the signs of these RG functions differ from those in, e.g., Ref. 1 since we consider the bare theory here.

Finally, we need the RG function ρ, Eq. (3.9c). We have performed a one-loop renormalization of the singlet propagator, Eq. (2.4b), and found[18]

$$\rho(x_0)=k_0+O(g_0^2)\ . \tag{3.13}$$

We see that k_0 is an irrelevant operator (its bare dimension is -1), and to one-loop order it has no anomalous dimension. While we do not know whether or not this result will hold to all orders, we note that an anomalous dimension of k_0 (i.e., a renormalization of the charge) would have profound consequences for the scaling behavior of the conductivity at the MIT. The DOS, on the other hand, would not be qualitatively affected. We can therefore restrict ourselves to the following scaling considerations. Let b be the cutoff dilatation factor ($\Lambda\to\Lambda/b$), and let the anomalous dimension of k_0 be $-1-x$, i.e., $k_0(b\to\infty)\sim b^{-1-x}$. x is determined by the renormalization of the charge in the screening wave number: The running charge will go as $e(b)\sim b^{x(D-1)/2}$. In a charged system, e cannot scale to zero, so $x\geq 0$.[19] x is related to a dynamical exponent. From perturbation theory, Eq. (2.7), we know that one of the combinations in which the external frequency occurs is $\Omega g_0\tilde{H}k_0^{1+\varepsilon}/\Lambda^D$. This defines a frequency scale and a dynamical exponent which was denoted by z_3 in Ref. 1, $\Omega\sim b^{-z_3}$ with $z_3=D-(1+\varepsilon)(1+x)$. The other two frequency scales, and their dynamical exponents (denoted by z_1 and z_2, respectively) are related to the combination $\Omega g_0\tilde{H}h_0/\Lambda^D$ in Eq. (2.7) and to a combination involving the spin-triplet interaction constant k_t,[1] which we have neglected. In terms of z_3 we can effectively write

$$\rho(x_0)=\frac{D-z_3}{D-1}k_0\ , \tag{3.14}$$

with $z_3=1+O(\varepsilon^2)\leq 1$.

The scaling behavior of the irrelevant operator k_0 determines a critical exponent, viz., z_3. However, this is of no consequence for the DOS since the time scale determined by z_3 is not the dominant one, as we will see. Consequently, all k_0 does is to produce corrections to scaling. These can be derived from including k_0 in the scaling functions, and the k_0 dependence of the RG functions we do not have to keep as well. We therefore replace Eq. (3.10a) by its values at $k_0=0$,

$$\gamma(x_0)=-g_0/2\varepsilon+O(g_0^3)\ . \tag{3.10′}$$

We close this section with four remarks. (1) In the case of the MI class, h_0 also flows to zero, cf. Eq. (3.11b). However, it does so much more slowly than k_0, and so $k_0^{1+\varepsilon}/h_0$ in Eq. (3.10a) still flows to zero. (2) In contrast to k_0, h_0 determines the leading critical time scale and the leading dynamical exponent z_1. It is therefore dangerously irrelevant[20] for the dynamics and must be treated with care. (3) A recalculation of the RG functions β and ζ might also reveal a dependence on k_0 (the results quoted above were obtained by effectively setting $k_0=0$). These dependencies on the irrelevant operator k_0 can be neglected for the same reason as in the case of γ. (4) If we had not introduced k_0 but instead kept the length scale $1/\kappa_D$, then the RG function γ, Eq. (3.10a), would depend on Λ through the combination Λ/κ_D. (This would be a particularly tempting option if z_3 was equal to 1, i.e., if κ_D was not renormalized.) This would be, of course, equivalent to solving the k_0-flow equation and inserting the result in the other RG functions. This is a, somewhat trivial, counterexample to the widespread conviction that RG functions must not depend on the scale. This erroneous belief seems to originate in the use of momentum-shell RG methods, which indeed do not allow for such a scale dependence. This question is relevant for the MIT problem in the presence of cooperons, and we will come back to it in the discussion in Sec. IV.

IV. SCALING PROPERTIES OF THE DENSITY OF STATES

A. Solution of the Callan-Symanzik equation

As a linear PDE of first order, the CS equation can be solved by the method of characteristics. For our purposes it is convenient to write the solution in the form[14]

$$\Gamma^{(1)}(\Omega;g_0,h_0,k_0;\Lambda)=\exp\left[-\tfrac{1}{2}\int_0^{\ln b}dl\,\gamma(g_0(e^l))\right]\Gamma^{(1)}(\Omega b^D,g_0(b),h_0(b),k_0(b);\Lambda)\ . \tag{4.1}$$

Here b is again the cutoff dilatation factor ($\Lambda\to\Lambda/b$), and the running coupling constants are given as solutions of the ordinary differential equations,

$$\frac{dg_0}{d\ln b}=-\beta(g_0(b))\ , \tag{4.2a}$$

$$\frac{dh_0}{d\ln b}=-\zeta(g_0(b),h_0(b),\gamma_t^0(b))\ , \tag{4.2b}$$

$$\frac{dk_0}{d\ln b}=\rho(k_0(b))\ , \tag{4.2c}$$

with β, ζ, and ρ given by Eqs. (3.11) and (3.12). In the case of the MF class, where ζ depends on γ_t^0, we need the additional flow equation, Eq. (3.12c),

$$\frac{d\gamma_t^0}{d\ln b}=\tfrac{1}{8}g_0(b)\{1-[\gamma_t^0(b)]^2\}+O(g_0^2)\ . \tag{4.2d}$$

The initial condition for Eqs. (4.2) is $x_0(b=1)=x_0$, with x_0 the set of physical parameters defined by Eqs. (3.7).

B. Asymptotic scaling, and corrections to scaling

Let us first derive the general asymptotic scaling properties of the DOS from Eqs. (4.1) and (4.2). The fixed points (FP) of the RG are given by those points x_0^* in parameter space for which the parameters do not charge under scaling. For our simple flow equations, x_0^* is given by the zeros of the RG functions, $\beta(x_0^*)=\zeta(x_0^*)=\cdots=0$. From Eqs. (3.11)–(3.13) we find for the MI class

$$(x_0^*)_{\rm MI}=(g_0^*,h_0^*,k_0^*)=(4\varepsilon,0,0)\ , \tag{4.3a}$$

and for the MF class

$$(x_0^*)_{\rm MF}=(g_0^*,h_0^*,(\gamma_t^0)^*,k_0^*)=\left[\frac{2\varepsilon}{1-\ln 2},\frac{h_0}{2}(1+\gamma_t^0),1,0\right]\ . \tag{4.3b}$$

In our case the FP value of g_0 coincides with the physical critical disorder g_0^c, $g_0^c=g_0^*$. We denote the deviation from the critical value by $t=1-g_0/g_0^c$. For g_0 close to g_0^c, and for values of b that are not too large, t grows as $t(b)=tb^{1/\nu}$ with $\nu=-1/\beta'(g_0^*)$ the correlation length exponent. The flow of h_0 is characterized by an exponent κ, $h_0(b)=h_0 b^{\kappa}$, and that of k_0 by the dynamical exponent z_3 as explained in connection with Eq. (3.14). Finally, the value of the DOS exponent β is determined by the FP value of the RG function γ: $\beta=-\nu\gamma(x_0^*)/2$. Asymptotically close to the critical point we can replace $\gamma(g_0(e^l))$ in Eq. (4.1) by $\gamma(g_0^*)$. We then find for the DOS a homogeneity or scaling relation,

$$N(\Omega;t,h_0,k_0;\Lambda)=b^{-\beta/\nu}N(\Omega b^D;tb^{1/\nu},h_0b^{\kappa},k_0b^{-(D-z_3)/(D-1)};\Lambda)\ , \tag{4.4}$$

with $\Omega=|\varepsilon-\varepsilon_F|$.

The exponents are easily obtained to one-loop order from Eqs. (3.10)–(3.12). We obtain

$$\nu=1/\varepsilon+O(1)\ , \tag{4.5}$$

for either universality class. The exponents β and κ are different for the two universality classes under consideration. One finds

$$\beta=\begin{cases}\dfrac{1}{\varepsilon}+O(1) & \text{(class MI)}\\[2mm] \dfrac{1}{2\varepsilon}\dfrac{1}{1-\ln 2}+O(1) & \text{(class MF)}\end{cases} \tag{4.6}$$

and

$$\kappa=\begin{cases}-\varepsilon/2+O(\varepsilon^2) & \text{(class MI)}\\ 0+O(\varepsilon^2) & \text{(class MF)}\ .\end{cases} \tag{4.7}$$

All of these values agree with Ref. 8. For z_3 we have

$$z_3=1+O(\varepsilon^2)\ , \tag{4.8}$$

for either universality class.

Let us now consider the disorder dependence of N at $\Omega=0$. Since N is dimensionless, at $\Omega=0$ it cannot depend on Λ. The k_0 dependence vanishes asymptotically. κ is either negative (class MI) or zero (class MF). By setting $b=t^{-\nu}$ we find

$$N(\Omega=0,t)\sim t^{\beta} \ . \tag{4.9a}$$

In order to find the corrections to this asymptotic scaling behavior we have to keep the irrelevant operators h_0 and k_0. From perturbation theory, Eq. (2.7), we know that at $\Omega=0$ they appear only in the combination $K_0=k_0^{1+\varepsilon}/h_0$. Further, for small values of K_0 the leading K_0 dependence is $K_0^{a(\varepsilon)}$ with

$$a(\varepsilon)=\frac{\varepsilon}{1-\varepsilon}\Theta(\tfrac{1}{2}-\varepsilon)+\Theta(\varepsilon-\tfrac{1}{2}) \ . \tag{4.9b}$$

This yields

$$N(\Omega=0,t)\sim t^{\beta}[1+\text{const}\times t^{a(\varepsilon)\nu(D-z_3+\kappa)}] \ . \tag{4.9c}$$

The term in the square brackets is the leading correction to scaling due to the irrelevant operators k_0 and h_0. Of course there are also corrections to scaling that are due to not being asymptotically close to the critical point, but have nothing to do with the irrelevant operators. These are analytic in t, and we have neglected them.

We now turn to the energy dependence of N at the critical disorder, $t=0$. Perturbation theory, Eq. (2.7), tells us that Ω appears in two combinations, viz., $\Omega_1=\Omega\tilde{H}g_0h_0/\Lambda^D$ and $\Omega_2=\Omega\tilde{H}g_0k_0^{1+\varepsilon}/\Lambda^D$. We thus find

$$N(\Omega;t=0)\sim b^{-\beta/\nu}N(\Omega b^{z_1},\Omega b^{z_3}) \ , \tag{4.10a}$$

where

$$z_1=D+\kappa \ . \tag{4.10b}$$

Since the two frequency arguments enter additively, the largest dynamical exponent is the dominant one. To one-loop order this is z_1. Setting $b=\Omega^{-1/z_1}$ we obtain

$$N(\Omega,t=0)\sim\Omega^{\beta/\nu z_1}[1+\text{const}\times\Omega^{a(\varepsilon)(1-z_3/z_1)}] \ . \tag{4.11}$$

Here Ω can either be interpreted as $|\varepsilon-\varepsilon_F|$ at zero temperature, or as temperature at $\varepsilon=\varepsilon_F$. The asymptotic scaling behavior is again in agreement with Ref. 8. To obtain the correction to scaling we have used the perturbative result, Eq. (2.7), that far from criticality (i.e., large Ω_1) and small Ω_2, N depends on its second frequency argument like $\Omega_2^{a(\varepsilon)}$. The leading correction to scaling in Eqs. (4.9) and (4.11) due to k_0 and h_0 is the same as one would obtain from keeping the k_0 and h_0 dependence of, e.g., the RG function $\gamma(x_0)$. This justifies the replacement of Eq. (3.10) by (3.10′).

C. Discussion

1. Relation to previous results

As we have mentioned before, our results for the asymptotic critical behavior of the DOS, Eqs. (4.9) and (4.11) with exponents given by Eqs. (4.5)–(4.7) and (4.10b), agree with those obtained by Finkel'stein in Ref. 8. They disagree with Ref. 9. The source of this disagreement is the behavior of the RG function $\gamma(x_0)$, Eq. (3.10a). As we have mentioned after Eq. (3.10b), in $D=2$ γ is a singular (viz., logarithmic) function of the irrelevant variable k_0. If we used, in the spirit of the ε expansion, this two-dimensional (2D) form of γ in Eq. (4.1), then we would recover the results of Ref. 9. Effectively, this procedure turns the $1/\varepsilon$ in the exponent β, Eq. (4.6), into a $\ln t$ or $\ln\Omega$. The principal problem with this procedure is that it ignores the fact that k_0 is dangerously irrelevant in $D=2$, but *not* in any $D>2$. Therefore using the 2D result in the FP description leads to a spuriously strong singularity in $D=2+\varepsilon$. Our treatment is in the spirit of Parisi:[21] We apply the RG directly in $D=2+\varepsilon$ dimensions rather than expanding everything about $D=2$. This avoids the problem mentioned above. Of course, in $D=2$ the dangerous irrelevancy of k_0 should have consequences, some of which are discussed in the next subsection.

2. Some remarks concerning $D=2$

In $D=2$ the system always scales towards strong coupling, i.e., large g_0, and there is no nontrivial FP and no phase transition. However, in many respects the system still shows critical behavior near $g_0=0$ just as, e.g., the 1D Ising model shows critical behavior near zero temperature. While the nature of the ground state is not known, we can still apply our RG methods to describe the remnants of critical behavior at intermediate scales before the crossover to strong-coupling behavior occurs. For $g_0\ll 1$, and scale factors b small enough such that $g_0(b)\cong g_0$, we can solve Eq. (4.1) to obtain

$$N(\Omega)\sim\exp\left[-\frac{g_0}{2}(\ln\Omega)^2\right] \ , \tag{4.12}$$

where Ω is either $|1-\varepsilon/\varepsilon_F|$ at $T=0$, or T/T_0 with some microscopic temperature scale T_0 at $\varepsilon=\varepsilon_F$. Equation (4.12) is valid in an intermediate energy or temperature region where Ω is larger than a scale $\Omega^*\sim e^{-1/g_0}$. For $\Omega\leq\Omega^*$ a crossover to some unknown behavior occurs.

3. General aspects of scaling and the structure of the RG

Let us conclude by discussing two general aspects of the RG method employed in this paper and putting them in the appropriate contexts.

The first point concerns the irrelevance of k_0, and possible scale dependence of the RG factors. As we already mentioned at the end of Sec. III, an alternative to introducing the irrelevant operator k_0 would have been to keep κ_D as an additional momentum scale. This would have led to a scale dependence of the RG function γ. It may be more natural, especially if the additional momentum has an anomalous dimension, to introduce an additional coupling constant as we have done, but the two descriptions are equivalent. One might think of situations where the structure of the theory is more complicated, and the coupling constants are harder to identify than in the present case. Under such circumstances, the field-theoretic RG would take any additional (irrelevant) operators that have not been identified explicitly into account by means of a scale dependence of the RG functions. By contrast, a Wilson-type momentum-shell RG, which does not allow for scale-dependent RG functions, would effectively replace the additional operators by their asymptotic values. In this way any corrections to scaling by the additional operators will be lost. We suspect that this is relevant for recent discussions concerning the MIT in the presence of cooperons. Reference 11 obtained logarithmic corrections to scaling from scale-dependent RG functions. While this is inconsistent with previous treatments of the problem by means of frequency-momentum-shell methods,[10] the above discussion could explain this discrepancy. If this was indeed the case, it should be possible to consistently introduce marginally irrelevant operators that produce the logarithmic corrections to scaling and eliminate the scale dependence of the flow equations. Efforts in this direction are under way.[12]

The second point concerns the existence of several time scales at the MIT. While their existence is already suggested by perturbation theory,[13] a full analysis of the consequences requires the RG. For the two universality classes studied here, we have two different time scales and associated dynamical exponents. A third exponent, z_2, which is related to the spin-triplet channel, is absent in the MI class and equal to z_1 in the MF class. *A priori* it is impossible to tell which time scale is the dominant one: Depending on the particular combination in which the various time scales appear in the scaling function, it is conceivable that either the largest or the smallest dynamical exponent is the dominant one. In the present case, one-loop perturbation theory suggests the former, i.e., z_1 is the dominant exponent and z_3 is related to an irrelevant time scale. Of course, the irrelevant time scale could be dangerous, but perturbation theory gives no indication of this to be the case. If at higher loop order it should turn out that $z_3 < 1$, then the corresponding time scale would indeed be dangerously irrelevant for the conductivity, but not for the DOS. The reason is that the conductivity is proportional to $e^2 \sim (\Lambda/k_0)^{(d-1)}$, and thus is a singular function of Λ/k_0 for $\Lambda/k_0 \to 0$. No such singular dependence is expected for the DOS, which depends on the charge only through the singlet propagator $\mathcal{D}^s$, Eq. (2.4b). At zero frequency the propagator is massless, and z_3 just determines how the soft propagator at zero frequency is approached. Since the corresponding time scale is not the dominant one, this does not affect the asymptotic scaling behavior, but manifests itself only in the corrections to scaling, Eqs. (4.11) and (4.9).

ACKNOWLEDGMENTS

This work was supported by the NSF under Grant Nos. DMR-92-09879 and DMR-92-17496.

---

[1]For a recent review, see, e.g., D. Belitz and T. R. Kirkpatrick, Rev. Mod. Phys. (to be published).

[2]F. Wegner, Z. Phys. B **44**, 9 (1981).

[3]B. L. Altshuler and A. G. Aronov, Solid State Commun. **30**, 115 (1979); B. L. Altshuler, A. G. Aronov, and P. A. Lee, Phys. Rev. Lett. **44**, 1288 (1980).

[4]A. L. Efros and B. I. Shklovskii, J. Phys. C **8**, L49 (1975).

[5]R. C. Dynes and J. P. Garno, Phys. Rev. Lett. **46**, 137 (1981); W. L. McMillan and J. Mochel, *ibid.* **46**, 556 (1981).

[6]K. B. Efetov, A. I. Larkin, and D. E. Khmelnitskii, Zh. Eksp. Teor. Fiz. **79**, 1120 (1980) [Sov. Phys. JETP **52**, 568 (1980)].

[7]F. Wegner, Z. Phys. B **35**, 207 (1979).

[8]A. M. Finkel'stein, Zh. Eksp. Teor. Fiz. **86**, 367 (1984) [Sov. Phys. JETP **59**, 212 (1984)].

[9]C. Castellani, C. Di Castro, P. A. Lee, and M. Ma, Phys. Rev. B **30**, 527 (1984).

[10]A. M. Finkel'stein, Z. Phys. B **56**, 189 (1984).

[11]T. R. Kirkpatrick and D. Belitz, Phys. Rev. Lett. **70**, 974 (1993).

[12]T. R. Kirkpatrick and D. Belitz (unpublished).

[13]T. R. Kirkpatrick and D. Belitz, Phys. Rev. B **41**, 11 082 (1990).

[14]See, e.g., M. Le Bellac, *Quantum and Statistical Field Theory* (Clarendon, Oxford, 1991).

[15]This is true only for the case of a long-ranged Coulomb interaction which we consider here. In the literature the case of a short-ranged model interaction has also been discussed, cf. Refs. 9 and 1.

[16]While renormalizability has never been proved, there is empirical evidence for it from perturbative calculations, cf. C. Castellani, C. Di Castro, and G. Forgacs, Phys. Rev. B **30**, 1593 (1984); C. Castellani, C. Di Castro, and S. Sorella, *ibid.* **34**, 1349 (1986); D. Belitz and T. R. Kirkpatrick, Nucl. Phys. **B316**, 509 (1989); and Ref. 13.

[17]We follow Ref. 14 in calling all partial differential equations of the type of Eq. (3.8) Callan-Symanzik equations.

[18]D. Belitz and T. R. Kirkpatrick (unpublished).

[19]M. P. A. Fisher, G.Grinstein, and S. M. Girvin, Phys. Rev. Lett. **64**, 587 (1990).

[20]M. E. Fisher, in *Advanced Course on Critical Phenomena,* edited by F. W. Hahne (Springer, Berlin, 1983), p. 1.

[21]G. Parisi, J. Stat. Phys. **23**, 49 (1980).