Landau Theory for a Metal-Insulator Transition

T. R. Kirkpatrick
Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742

D. Belitz
Department of Physics and Materials Science Institute, University of Oregon, Eugene, Oregon 97403
(Received 15 February 1994)

The nonlinear σ model for disordered interacting electrons is studied in spatial dimensions $d > 4$. The critical behavior at the metal-insulator transition is determined exactly, and found to be that of a standard Landau-Ginzburg-Wilson ϕ^4 theory with the single-particle density of states as the order parameter. All static exponents have their mean-field values, and the dynamical exponent $z = 3$. $\partial n/\partial\mu$ is critical with an exponent of 1/2, and the electrical conductivity vanishes with an exponent $s = 1$. The transition is qualitatively different from the one found in the same model in a $2 + \epsilon$ expansion.

PACS numbers: 71.30.+h, 64.60.Ak

Despite much progress in our understanding of the metal-insulator transition (MIT) in recent years [1], the search for a simple order parameter description of this phase transition in the spirit of a mean-field or Landau theory has so far proven futile. Early attempts at a mean-field theory of the Anderson transition of noninteracting electrons [2] failed because the most obvious simple order parameter (OP), viz., the single-particle density of states (DOS) at the Fermi level, turned out to be uncritical [3]. As a result, the Anderson transition can be described only in terms of a complicated functional OP [4]. At the Anderson-Mott transition of disordered interacting electrons, on the other hand, the DOS is generally believed to be critical, but it is not obvious how to construct an OP description in terms of it. Indeed, our understanding of this transition is largely based on a generalized matrix nonlinear σ (NLσ) model [5], for which no mean-field type fixed point (FP) is known, and which has been analyzed in terms of a small disorder expansion near $d = 2$. As a result, there is no simple description of the MIT analogous to, say, Weiss theory of ferromagnetism. This is all the more remarkable because of the technical similarity, first noted by Wegner [5], between the NLσ model description of the MIT and that of Heisenberg ferromagnets.

In this Letter we show that the NLσ model for interacting disordered electrons [5] possesses a saddle point solution which has all of the characteristic features of a Landau theory with the DOS as the OP. We further show that this solution corresponds to a renormalization group (RG) FP which is stable for $d > 4$. This establishes the exact (for $d > 4$) critical behavior at the MIT in this model, which can be summarized as follows: Let t be the dimensionless distance from the critical point at temperature $T = 0$, Ω the energy distance from the Fermi level, and Q the DOS. Q vanishes according to

$$Q(t, \Omega = 0) \sim t^{\beta}, \quad Q(t = 0, \Omega) \sim \Omega^{\beta/\nu z}. \qquad (1a)$$

For the critical exponents β, ν, and z, and for the exponents γ and η characterizing the OP susceptibility, we find

$$\nu = \beta = 1/2, \quad \gamma = 1, \quad \eta = 0, \quad z = 3. \qquad (1b)$$

All thermodynamic susceptibilities show the same critical behavior,

$$\chi(t, \Omega = 0) \sim t^{1/2}, \quad \chi(t = 0, \Omega) \sim \Omega^{1/3}. \qquad (2)$$

Here χ can stand for the critical parts of the density susceptibility $\partial n/\partial\mu$, the specific heat coefficient $\gamma = \lim_{T\to 0} C_V(t,T)/T$, or the spin suceptibility χ_s. In general all of these susceptibilities can also have an additive noncritical contribution which is an analytic function of t and Ω. As an argument of susceptibilities, Ω denotes the external frequency, and Ω and T can be used interchangeably in a scaling sense. The charge diffusion coefficient behaves like the OP, Eq. (1a). Assuming that $\partial n/\partial\mu$ has no noncritical background contribution, the electrical conductivity σ vanishes according to

$$\sigma(t, \Omega = 0) \sim t, \quad \sigma(t = 0, \Omega) \sim \Omega^{2/3}, \qquad (3)$$

so the conductivity exponent $s = 1$.

In what follows we first derive Eqs. (1) by explicitly constructing the saddle point solution, and then using RG techniques to show that it is stable for $d > 4$. We then use scaling arguments to obtain additional information about susceptibilities, which leads to Eqs. (2) and (3). We consider the matrix NLσ model of Refs. [1,5], i.e., a Gaussian field theory for a Hermitian matrix field $\tilde{Q}(\mathbf{x})$ with constraints $[\tilde{Q}(\mathbf{x})]^2 = \mathbb{1}$, with $\mathbb{1}$ the unit matrix, and $\mathrm{tr}\tilde{Q}(\mathbf{x}) = 0$. $\tilde{Q}$ is a classical field comprising two fermionic fields. It carries two Matsubara frequency indices n, m and two replica indices α, β (quenched disorder has been incorporated by means of the replica trick). The matrix elements $\tilde{Q}_{nm}^{\alpha\beta}(\mathbf{x})$ are in general spin quaternions, with the quarternion degrees of freedom describing the particle-hole and particle-particle channel, respectively. The action [1,5] can be written

0031-9007/94/73(6)/862(4)$06.00

$$S[\tilde{Q},\Lambda] = \frac{-1}{2G}\int d\mathbf{x}\ \mathrm{tr}\{\Lambda(\mathbf{x})[\tilde{Q}^2(\mathbf{x}) - \mathbb{1}] + [\partial_{\mathbf{x}}\tilde{Q}(\mathbf{x})]^2\} + 2H\int d\mathbf{x}\ \mathrm{tr}[\Omega\tilde{Q}(\mathbf{x})] - \frac{\pi T}{4}\sum_{u=s,t,c} K_u[\tilde{Q}(\mathbf{x})\circ\tilde{Q}(\mathbf{x})]_u. \tag{4}$$

Here G is a measure of the disorder, Ω is a diagonal matrix whose elements are the external Matsubara frequencies ω_n, and H is proportional to the free electron DOS. $K_{s,t,c}$ are coupling constants describing the electron-electron interaction in the particle-hole spin singlet, particle-hole spin triplet, and particle-particle or Cooper channel, respectively. $K_s < 0$ for repulsive interactions. $[\tilde{Q}\circ\tilde{Q}]_{s,t,c}$ denotes a product in frequency space which is given explicitly in Refs. [1,5]. Notice that we have enforced the constraint $\tilde{Q}^2 = \mathbb{1}$ by means of an auxiliary matrix field $\Lambda(\mathbf{x})$. The constraint $\mathrm{tr}\tilde{Q} = 0$ will be enforced explicitly at every stage of the theory.

The model, Eq. (4), is an effective one which was derived to capture the essence of the low-lying excitations of the system, i.e., the low-frequency large-wavelength behavior. The soft (i.e., diffusive) modes are given by correlation functions of the $\tilde{Q}_{nm}$ with $nm < 0$, while the DOS is determined by the average of $\tilde{Q}^{\alpha\alpha}_{nn}$ [1,5]. It is therefore convenient to separate $\tilde{Q}$ into blocks:

$$\tilde{Q}^{\alpha\beta}_{nm} = \Theta(nm)Q^{\alpha\beta}_{nm}(\mathbf{x}) + \Theta(n)\Theta(-m)q^{\alpha\beta}_{nm}(\mathbf{x}) + \Theta(-n)\Theta(m)(q^{+})^{\alpha\beta}_{nm}(\mathbf{x}). \tag{5}$$

The conventional treatment of the NLσ model [1,5] proceeds by integrating out $\Lambda(\mathbf{x})$, using the constraint $\tilde{Q}^2 = \mathbb{1}$ to eliminate Q, and expanding the action in powers of q. Here we use a different approach inspired by treatments of the $O(N)$ NLσ model in the large-N limit [6].

Since the further development will be closely analogous to the treatment of an $O(N)$ Heisenberg model, let us pause to point out the similarities between these models. The matrix elements Q_{nm} correspond to the massive σ component of the $O(N)$ vector field, while the q_{nm} correspond to the massless π fields. The disorder G plays the role of the temperature in the magnetic model, it is the control parameter for the phase transition. $H\Omega$ is in some respects analogous to the magnetic field conjugate to the order parameter σ. The last term in Eq. (4) has no analogy in the Heisenberg model. We will see that in the mean-field treatment of the matrix model presented here it plays a rather trivial, although crucial, role.

We proceed by integrating out the massless q field. Since the action, Eq. (4), is quadratic in $\tilde{Q}$ and hence in q, this can be done exactly. We then look for a saddle point of the resulting effective action $S[Q,\Lambda]$. This task is simplified by restricting ourselves to saddle point solutions that are spatially constant, diagonal matrices with diagonal elements Q^{α}_{n}, Λ^{α}_{n}. This *ansatz* is motivated by the fact that $\langle Q^{\alpha\beta}_{nm}\rangle$ has these properties. For simplicity we restrict ourselves to the particle-hole spin-singlet channel, i.e., we put $K_t = K_c = 0$. We will see later that this restriction does not influence the critical behavior. This last property is expected in a mean-field theory where all universality classes typically have the same critical behavior. We further use a short-range model interaction, so that K_s is simply a number. Again it can be shown that a Coulomb interaction leads to the same critical behavior [7]. We find for the saddle point solution,

$$(Q^{\alpha}_{n})^2 = 1 + \frac{G}{4}\sum_{\mathbf{p}}\sum_{m=-1}^{-\infty}\frac{2\pi TGK_s}{[p^2+\frac{1}{2}(\Lambda^{\alpha}_{n}+\Lambda^{\alpha}_{m})]^2}\left[1+\sum_{n_1=0}^{n-m-1}\frac{2\pi TGK_s}{p^2+\frac{1}{2}(\Lambda^{\alpha}_{n_1}+\Lambda^{\alpha}_{n_1-n+m})}\right]^{-1}, \tag{6a}$$

$$\Lambda^{\alpha}_{n} = 2GH\omega_n/Q^{\alpha}_{n}. \tag{6b}$$

We discuss several aspects of this result. First, $Q^{\alpha}_{n=0} \equiv Q$, which is the DOS normalized by the free electron result, decreases with increasing disorder (remember $K_s < 0$). This is the well known "Coulomb anomaly" in the DOS [8], and it is proportional to K_s as well as G. Technically, this is due to the replica structure of the theory: All terms on the right hand side of Eq. (6a) that are independent of K_s vanish in the replica limit and have not been shown. Physically, it reflects the fact that the Coulomb anomaly is due to the *interplay* between interactions and disorder. It is important that, in the light of Refs. [2,3], our method to construct a mean-field theory had better *not* work in the absence of interactions (i.e., for $K_s = 0$). This expectation is borne out explicitly by Eq. (6a), which correctly yields $Q \equiv 1$ for $K_s = 0$. Second, Eqs. (6) constitute an integral equation for Q^{α}_{n} which involves an integration over all frequencies. As noted above, the NLσ model is designed to describe only low-frequency behavior, and cannot be trusted at high frequencies. However, on physical grounds it is clear that Q^{α}_{n} tends to a constant at large frequencies, so for fixed K_s Q will vanish at a critical value G_c of G. We can easily determine the critical behavior for $\delta G \equiv G - G_c < 0$: We find $Q \sim (-\delta G)^{1/2}$, which is the first relation in Eq. (1a). Furthermore, the leading low-frequency behavior for $G = G_c$ involves only integrals up to the external frequency, which is the region where the theory is controlled. For $d > 4$ we find $Q(\Omega) \sim \Omega^{1/3}$, which is the second relation in Eq. (1a) with the exponent values given in Eq. (1b).

We have shown that our saddle point actually corresponds to a minimum of the free energy by expanding the action to second order in the fluctuations $\delta Q, \delta\Lambda$ about the saddle point. Details of this calculation will be reported elsewhere [7]. The result is a positive definite Gaussian matrix, so the saddle point is indeed a minimum. One can then integrate out $\delta\Lambda$ to obtain the order parameter correlation function $G_{n_1 n_2, n_2 n_4}(\mathbf{k})$. At zero frequency, and close to the critical point (i.e., for $Q \to 0$) the result simplifies substantially and we find

$$G_{00,00}(\mathbf{k}) = \frac{G/4}{k^2 + Q^2/\xi_0^2}, \tag{7}$$

where the bare correlation length ξ_0 is given in terms of a complicated integral which is finite for $d > 4$. From Eq. (7) we obtain three more critical exponents: $\nu = \beta = 1/2, \eta = 0$, and $\gamma = 2\nu = 1$, cf. Eq. (1b).

Apart from the DOS we are interested in the transport properties. Let us first consider the charge diffusion coefficient D_c. As a hydrodynamic quantity it can be obtained from the NLσ model by a direct calculation of the particle-hole spin-singlet q propagator [1,5]. Inserting the saddle point solution, Eqs. (6), into Eq. (4), one reads off the q-vertex function, and a matrix inversion yields the corresponding propagator. The latter has the usual diffusion pole structure, and D_c is obtained as the coefficient of the momentum squared. We find

$$D_c = \frac{Q}{G(H + K_s Q)} \ . \tag{8}$$

We see that at the transition D_c vanishes like the OP. The behavior of the conductivity can be obtained by multiplying D_c with $\partial n/\partial\mu$. The latter is less straightforward to obtain, since as a thermodynamic quantity it is not simply given by the hydrodynamic behavior. We will determine its critical behavior from Eqs. (11) below.

We now use RG techniques to show that these results represent the *exact* critical behavior of the model in $d > 4$. The RG will also provide an easier route to a derivation of Eqs. (2) and (3) than a direct calculation would be. Let us return to the action, Eq. (4). If we integrate out q, we get Eq. (4) with $\tilde{Q}$ replaced by Q plus terms obtained by contracting q fields. Of the latter, one term is linear in Λ. The resulting action can be written in the form

$$S[Q,\Lambda] = -c\int d\mathbf{x}\, \mathrm{tr}[\partial_{\mathbf{x}} Q(\mathbf{x})]^2 - \int d\mathbf{x}\, \mathrm{tr}[t\Lambda(\mathbf{x})] + 2H\int d\mathbf{x}\, \mathrm{tr}[\Omega Q(\mathbf{x})] - \frac{\pi T}{4}\sum_{u=s,t,c} K_u [Q(\mathbf{x}) \circ Q(\mathbf{x})]_u$$
$$+ u_1 \int d\mathbf{x}\, \mathrm{tr}[\Lambda(\mathbf{x}) Q^2(\mathbf{x})] + \int d\mathbf{x} \sum_{I,J} \Lambda_I(\mathbf{x}) (u_2)_{IJ} \Lambda_J(\mathbf{x}) + (\text{other terms}). \tag{9}$$

Here $c = u_1 = 1/2G$ and t is a matrix composed of $-(1/2G)\mathbb{1}$ and the term of $O(\Lambda)$ coming from the q contractions. $I \equiv (nm, \alpha\beta, i)$ where i labels the spin-quaternion space, and u_2 is a matrix which is finite for $d > 4$. The explicit expressions for t and u_2 will not be needed for our present purposes, they will be given elsewhere [7]. The "other terms" in Eq. (9) all come from contracting q fields. They can easily be constructed diagrammatically, but will turn out to be irrelevant for our purpose.

We now apply standard power counting to the action, Eq. (9). Our parameter space is spanned by $\mu = \{c, t, H, K_{s,t,c}, u_1, u_2\}$ and the coupling constants of the other terms. In looking for a RG fixed point, we follow Refs. [9] and [6] in fixing the exact scale dimensions of our fields Q and Λ to be $[Q] = d/2 - 1$, $[\Lambda] = d - 2$ (we define the scale dimension of a length to be -1). This corresponds to fixing the exponents η and $\tilde{\eta}$ defined by the wave number dependence of the Q- and Λ-correlation functions to be $\eta = 0$ and $\tilde{\eta} = d - 4$, respectively. With $[\Omega] = [T] = d$, we find the scale dimensions of the various coupling in Eq. (9) to be $[c] = 0, [t] = 2, [H] = 1 - d/2, [K_{s,t,c}] = 2 - d, [u_1] = [u_2] = 4 - d$. A power counting analysis of all of the other terms in Eq. (9) is straightforward. The result is that the coupling constants of all of these terms have scale dimensions which are smaller than $4 - d$ for $d > 4$ [10]. We conclude that the Gaussian FP given by $\mu^* = \{c, 0, 0, 0, \ldots\}$ is stable for $d > 4$. The only relevant parameter is t, and the correlation length exponent ν is $\nu = 1/[t] = 1/2$. c is marginal, as expected, and all other parameters, including H and $K_{s,t,c,}$, are irrelevant (the combinations $H\Omega$ and $K_{s,t,c}T$ are, of course, relevant, which reflects the fact that a finite frequency or wave number drives the system away from the critical point). This Gaussian FP obviously corresponds to the saddle point solution discussed above. The stability of the FP proves that the critical behavior obtained from the saddle point solution is exact in $d > 4$ [10].

The RG arguments given above imply that the order parameter obeys the scaling relation

$$Q(t, H\Omega, u_1, u_2, \ldots) = b^{1-d/2} Q(t b^{1/\nu}, H\Omega b^{1+d/2}, u_1 b^{4-d}, u_2 b^{4-d}, \ldots), \tag{10}$$

where $\nu = 1/2$, and we have suppressed all parameters with scale dimensions smaller than $4 - d$. The exponents β and z follow from Eq. (10) by using standard arguments [9]. The crucial point is that u_1 and u_2 are dangerous irrelevant variables: Solving the theory explicitly in the saddle point approximation, we see that $Q(t, \Omega = 0, u_1 \to 0) \sim u_1^{-1/2}$, and $Q(t = 0, \Omega, u_1 \to 0, u_2 \to 0) \sim (u_2/u_1^2)^{1/3}$. Equation (10) then immediately yields $\beta = 1/2, z = 3$ in agreement with Eq. (1b). We emphasize that K_s is *not* dangerously irrelevant, even though $Q(t = 0, \Omega)$ is proportional to $\sqrt{K_s}$. The point is that the vanishing of the (bare) K_s just shifts the transition point to infinity, making the mean-field transition inaccessible as it should be at $K_s = 0$ [2,3], but does not change the critical behavior.

We finally derive Eqs. (2) and (3). To this end we notice that all susceptibilities have been identified in terms of the coupling constants of the NLσ model: $\partial n/\partial\mu \sim H + K_s, \chi_s \sim H + K_t, \gamma \sim H$ [1,11]. H, K_s, and K_t are all irrelevant, which means that all three susceptibilities vanish at the critical point. Furthermore, $K_{s,t}$ scale to zero faster than H, so the asymptotic scaling behavior of all three susceptibilities is the same and given by that of H, or γ. The only remaining difficulty is to correctly incorporate the dangerous irrelevant variables. Let us consider the singular part f_s of the free energy density f, which satisfies the scaling equation [1],

$$f_s(t,T,u_1,\ldots) = b^{-(d+z)} f_s(tb^{1/\nu}, Tb^z, u_1 b^{4-d}, \ldots). \tag{11a}$$

In the critical regime we have $f_s \sim Q^2 \sim 1/u_1$, and hence

$$f_s(t,T) = b^{-(4+z)} f_s(tb^{1/\nu}, Tb^z). \tag{11b}$$

We see that hyperscaling breaks down in the usual way: d in Eq. (11a) gets replaced by 4. By differentiating Eq. (11b) twice with respect to temperature, and using $z = 3$, we find

$$\gamma_s(t,T) = b^{-1}\gamma_s(tb^2, Tb^3), \tag{12}$$

and hence Eqs. (2) [12]. Finally, combination of this result with Eq. (8) yields Eq. (3). We note that the criticality of $\partial n/\partial\mu$ is very remarkable, since it does not show in the $2+\epsilon$ expansion treatment of the same model [1,5], while it is an important feature of the Mott-Hubbard transition, which recently has enjoyed a revival of interest [13].

Several questions arise from these results: (1) What is the connection between the present results and those obtained in $d = 2 + \epsilon$, where $\partial n/\partial\mu$ is found to be uncritical in perturbation theory [14]? A possible answer is that $\partial n/\partial\mu$ has an essential singularity near $d = 2$ which is invisible in the $2 + \epsilon$ expansion. Another possibility is that the Gaussian FP discussed here is not continuously connected to the one found near $d = 2$. (2) Can the present techniques be extended to the more general field theory [1] underlying the NLσ model? If that is the case, one would expect the upper critical dimension of the resulting OP description to be 6, or, for technical reasons similar to those in Ref. [2], 8. Which of the two models would capture more of the physics relevant in $d = 3$ would *a priori* be unclear. These problems will be pursued in the future [7].

This work was supported by the NSF under Grants No. DMR-92-17496 and No. DMR-92-09879.

[1] For a recent review, see, e.g., D. Belitz and T. R. Kirkpatrick, Rev. Mod. Phys. **66**, 261 (1994).

[2] A. B. Harris and T. C. Lubensky, Phys. Rev. B **23**, 2640 (1981); see also J. B. Marston and I. Affleck, Nucl. Phys. **B290** **[FS20]**, 137 (1987).

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

[4] Y. V. Fyodorov and A. D. Mirlin, Phys. Rev. Lett. **67**, 2049 (1991).

[5] A. M. Finkel'stein, Zh. Eksp. Teor. Fiz. **84**, 168 (1983) [Sov. Phys. JETP **57**, 97 (1983)]; this is a generalization of Wegner's model for noninteracting electrons, F. Wegner, Z. Phys. B **35**, 207 (1979).

[6] See, e.g., J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena* (Clarendon, Oxford, 1989).

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

[8] B. L. Altshuler and A. G. Aronov, Solid State Commun. **30**, 115 (1979).

[9] S.-K. Ma, *Modern Theory of Critical Phenomena* (Benjamin, Reading, MA, 1976); M. E. Fisher, in *Advanced Course on Critical Phenomena*, edited by F. W. Hahne (Springer, Berlin, 1983), p. 1.

[10] This conclusion is based on power counting at zero-loop order. One should keep in mind that terms generated by the RG at higher loop order can in principle raise the upper critical dimension d_c^+; for an example, see Ref. [2]. If this should happen here as well, all of our results would still hold, but only for $d > d_c^+ > 4$ instead of $d > 4$.

[11] C. Castellani, C. DiCastro, P. A. Lee, M. Ma, S. Sorella, and E. Tabet, Phys. Rev. B **33**, 6169 (1986).

[12] We note that Eqs. (11) by themselves give information only about the singular or critical part of γ. This is why Eqs. (2) in general hold only for the critical parts of the susceptibilities. If one uses in addition the information that H scales to zero, then one concludes that the full susceptibilites, not just their critical parts, vanish. However, more complete models might contain noncritical background contributions that are absent from our effective model which concentrates on the critical behavior. Of course this would not change the critical behavior of anything except possibly the conductivity: If, due to physical processes not contained in our model, there was a noncritical background contribution to $\partial n/\partial\mu$, then the critical behavior of the conductivity would be the same as that of the charge diffusion coefficient. Alternatively, noncritical processes not included in our model might lead to bare interaction constants in Eq. (4) that diverge at $T = 0$. Such terms at most would cause some of the soft modes in our model to be absent. For $d > 4$ this would not modify the critical behavior.

[13] For a recent review, see, D. Vollhardt, in *Correlated Electron Systems*, edited by V. Emery (World Scientific, Singapore, 1993).

[14] In $d = 2$ one does not expect logarithmic corrections to $\partial n/\partial\mu$ to all orders in perturbation theory. The reason is that $\partial n/\partial\mu$ is given as a frequency integral over a quantity which itself is only logarithmically singular.