The Anderson-Mott 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, and Department of Physics, University of Maryland, College Park, Maryland 20742

The interacting disordered electron problem is reviewed with emphasis on the quantum phase transitions that occur in a model system and on the field-theoretic methods used to describe them. An elementary discussion of conservation laws and diffusive dynamics is followed by a detailed derivation of the extended nonlinear sigma model, which serves as an effective field theory for the problem. A general scaling theory of metal-insulator and related transitions is developed, and explicit renormalization-group calculations for the various universality classes are reviewed and compared with experimental results. A discussion of pertinent physical ideas and phenomenological approaches to the metal-insulator transition not contained in the sigma-model approach is given, and phase-transition aspects of related problems, like disordered superconductors and the quantum Hall effect, are discussed. The review concludes with a list of open problems.

CONTENTS

 0034-6861/94/66(2)/261(120)/$17.00

I. INTRODUCTION

In the field of continuous phase transitions, metal-insulator transitions play a special role. In the first place, they are not nearly so well understood, either experimentally or theoretically, as the classic examples of liquid-gas critical point, Curie point, λ point in ^{4}He, etc. In the second place, one important subclass of metal-insulator transition consists of what are now called quantum phase transitions, i.e., continuous phase transitions that occur at zero temperature, so that the fluctuations determining the critical behavior are quantum mechanical rather than thermal in nature.

Metal-insulator transitions can be divided into two categories (see, for example, Mott, 1990). In the first category, some change in the ionic lattice, such as a structural phase transition, leads to a splitting of the electronic conduction band and hence to a metal-insulator transition. In the second category the transition is purely electronic in origin and can be described by models in which the lattice is either fixed or altogether absent as in models of the "jellium" type. It is this second category which forms the subject of the present article. Historically, the second category has again been divided into two classes, one in which the transition is triggered by electronic correlations and one in which it is triggered by disorder. The first case is known as a Mott or Mott-Hubbard transition, the second as an Anderson transition.

Mott's original idea (Mott, 1949, 1990) of the correlation-induced transition was intended to explain why certain materials with one electron per unit cell, e.g., NiO, are insulators. Mott imagined a crystalline array of atomic potentials with one electron per atom and a Coulomb interaction between the electrons. For sufficiently small lattice spacing, or high electron density, the ion cores will be screened, and the system will be metallic. Mott argued that for lattice spacing larger than a critical value the screening will break down, and the system will undergo a *first-order* transition to an insulator. This argument depended on the long-range nature of the Coulomb interaction. A related, albeit continuous, metal-insulator transition is believed to occur in a tight-binding model with a short-ranged electron-electron interaction known as the Hubbard model (Anderson, 1959a; Gutzwiller, 1963; Hubbard, 1963). The model Hamiltonian is

$$\hat{H}=t\sum_{\langle i,j\rangle}(\hat{a}_{i\uparrow}^{+}\hat{a}_{j\uparrow}+\hat{a}_{i\downarrow}^{+}\hat{a}_{j\downarrow})+U\sum_{i}\hat{n}_{i\uparrow}\hat{n}_{i\downarrow}\ , \tag{1.1a}$$

where $\hat{a}_{i\sigma}^{+}$ and $\hat{a}_{i\sigma}$ are creation and annihilation operators, respectively, for electrons with spin σ at site i, the summation in the tight-binding term is over nearest neighbors, and

$$\hat{n}_{i\sigma}=\hat{a}_{i\sigma}^{+}\hat{a}_{i\sigma}\ . \tag{1.1b}$$

Here t is the hopping matrix element, and U is an on-site repulsion energy ($U>0$). Despite its simplicity, remarkably little is known about the Hubbard model. In one dimension ($1-d$) it has been solved exactly by Lieb and Wu (1968), who showed that at half filling the ground state is an antiferromagnetic insulator for any $U>0$. In higher dimensions, various approximations have suggested that the model with one electron per site shows a continuous metal-insulator transition at zero temperature as a function of U (see Mott, 1990). This is generally believed to be true, but has not been firmly established. Recently the Hubbard model has received renewed atten-

tion, mostly because models closely related to it have been suggested to contain the relevant physics for an understanding of high-T_c superconductivity (Anderson, 1987). This renewed interest has spawned new approximation schemes as well as new rigorous results. For instance, it has been proven (among other things) that the ground state of the Hamiltonian, Eqs. (1.1), at half filling has spin zero in all dimensions (Lieb, 1989). The Hubbard model has also been studied in the limit of infinite dimension (Metzner and Vollhardt, 1989). It turns out that in this limit the model can be reduced to a one-dimensional problem that has been solved numerically. A number of nontrivial results have been obtained. In particular, it has been established that in $d=\infty$ with increasing U there is a Mott transition from a metallic to an insulating state (Georges and Krauth, 1992; Jarrel, 1992; Rozenberg *et al.*, 1992).

Much more is known about the Anderson transition (Anderson, 1958; for a recent review see Lee and Ramakrishnan, 1985), which can be described by the Anderson model,

$$\hat{H}=t\sum_{\langle i,j\rangle}\hat{a}_i^{+}\hat{a}_j+\sum_i \varepsilon_i\hat{a}_i^{+}\hat{a}_i \ . \tag{1.2}$$

This is a model for noninteracting electrons, so spin provides only trivial factors of two and can be omitted. The ε_i are randomly distributed site energies, governed by some distribution function characterized by a width W. Anderson argued that for W/t large but finite the system is an insulator. In $d=1$ it actually is insulating for all $W>0$ (Mott and Twose, 1961; Borland, 1963). For large W, or near the band tails, the states were proven to be localized (Fröhlich *et al.*, 1985). In $d=3$, Anderson predicted a metal-insulator transition to occur at a nonzero value of W/t, a prediction which was confirmed numerically (Schönhammer and Brenig, 1973). This as well as all other early (pre-1979) work, both numerical and analytical, also found a metal-insulator transition in $d=2$, which was later shown to be incorrect (see below).

The reason for the insulating behavior at large W is that the electrons become "trapped" or "localized" in the potential fluctuations. This is not an intrinsically quantum-mechanical phenomenon. Anderson's work had been motivated in part by work on classical percolation (Broadbent and Hammersley, 1957), and the classical Lorentz model (see, for example, Hauge, 1974) shows a transition between a diffusive and a localized phase much like the Anderson model (see Fig. 1). In the Lorentz model, a classical pointlike particle moves in a random array of fixed scatterers, often taken to be hard disks ($d=2$) or hard spheres ($d=3$). The localization does not require attractive potentials, but rather comes about by trapping of the particle in cages at sufficiently high scatterer density. If the scattering potentials are soft, the delocalization can be achieved not only by decreasing the density of scatterers, but also by increasing the energy of the scattered particle. An analogous effect exists in the quantum case: if W is not so large as to localize the whole band, then energies E_c separate localized states in the band tails from extended ones in the band center, and the metal-insulator transition can be triggered by sweeping the Fermi energy across the "mobility edge" E_c (Mott, 1966, 1990); see Fig. 2.

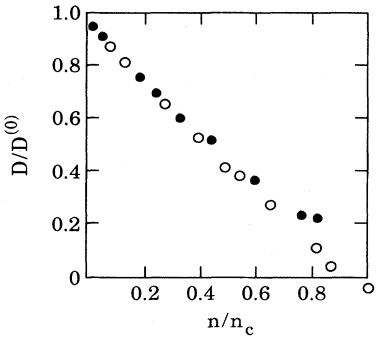

FIG. 1. Numerical simulation data for the diffusion coefficient D vs the scatterer density n of a 2-d Lorentz model: ●, Bruin (1972, 1974, 1978); ○, Alder and Alley (1978), and Alley (1979), as quoted by Götze *et al.* (1982). D is normalized by its Boltzmann value $D^{(0)}$, and n by its critical value n_c. After Götze *et al.* (1982).

A further important development occurred when Wegner (1976a) used real-space renormalization-group methods to argue that the dynamical conductivity could be written in the scaling form

$$\sigma(t,\Omega)=b^{-(d-2)}f(tb^{1/\nu},\Omega b^{d}) \ . \tag{1.3}$$

Here t is some dimensionless distance from the critical point (e.g., $t=|W-W_c|/W_c$ or $t=|E-E_c|/E_c$), Ω is the frequency, b is an arbitrary scale parameter, f is an unknown scaling function, and ν an unknown correlation length exponent. Equation (1.3) predicts that the static conductivity vanishes at the metal-insulator transition with an exponent $s=\nu(d-2)$, and that the dynamical

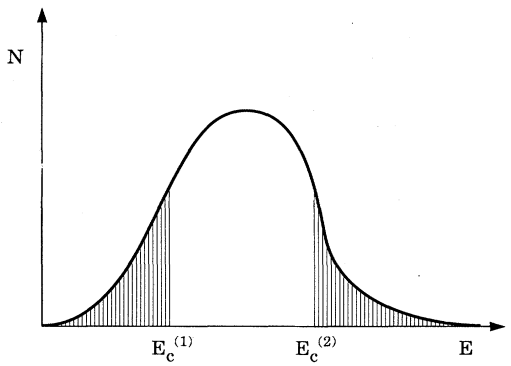

FIG. 2. Schematic picture of the density of states N vs the energy E in the Anderson model. $E_c^{(1,2)}$ are mobility edges, and the states in the shaded regions are localized.

conductivity at the critical point, $t=0$, goes[1] like $\Omega^{(d-2)/d}$. The importance of this result lay in the demonstration that the metal-insulator transition could be discussed in the canonical terms of critical phenomenon theory (see, for example, Ma, 1976; Fisher, 1983). This line of approach to the model was taken one step further with Wegner's (1979) mapping of the Anderson localization problem onto an effective field theory. The methods employed by Wegner and his followers are the central theme of this review and will be explained in detail in Sec. III.

Abrahams *et al.* (1979) conjectured that in $d=2$ all states are localized by arbitrarily weak disorder, as they are in $d=1$. In contrast to Anderson localization *per se,* this is a pure quantum effect, which is not present in the Lorentz model and which can be understood as an interference phenomenon (Bergmann, 1983, 1984). In contrast to the 1-d case, the effect for weak disorder in $d=2$ is only logarithmic for temperatures that are not too low. A key prediction was that in thin metallic films the resistance at moderately low temperatures should logarithmically increase with decreasing temperature, a prediction that has become known as the "weak-localization" effect. The observation of such logarithmic rises (Dolan and Osheroff, 1979; see also Bergmann, 1984) was widely hailed as a confirmation of the theory. Later it became clear that at least in some cases the agreement had been fortuitous, as the importance of interaction effects, which are neglected in the weak-localization model, was not appreciated early on. For many subsequent years the field enjoyed considerable activity, which has been reviewed by Bergmann (1984) and by Lee and Ramakrishnan (1985).

Shortly after the prediction of the weak-localization effect it was realized that very similar effects can be caused by a completely different physical mechanism. Altshuler and Aronov (1979a, 1979b) showed that the electron-electron interaction in 3-d weakly disordered systems leads to a square-root cusp in the tunneling density of states and to corresponding square-root anomalies in the temperature and frequency dependence of the specific-heat coefficient and the conductivity. The latter anomaly has the same functional form as the 3-d weak-localization contribution (Gorkov *et al.*, 1979). In 2-d the corresponding effects are logarithmic and, in particular, the conductivity was predicted to have a logarithmic temperature dependence just like the weak-localization effect, even if the interference effects that cause the latter are neglected (Altshuler, Aronov, and Lee, 1980; Fukuyama, 1980). It thus became clear that an observed anomaly in the conductivity by itself could not be taken as evidence for the presence of the weak-localization effect. Rather, one must try to separate the two effects by means of the above-mentioned thermodynamic and density-of-states anomalies, which accompany the interaction effect but not weak localization, or by means of the magnetoresistance. The magnetoresistance is negative for the weak-localization model, since a magnetic field destroys the phase coherence that is essential for producing the interference effect (Altshuler, Khmelnitskii, Larkin, and Lee, 1980), while it is zero or positive for various interaction models (Fukuyama, 1980; Altshuler *et al.*, 1981).

These perturbative considerations in the weak-disorder regime raised the question of whether and how interaction anomalies affect the metal-insulator transition. Since the Coulomb interaction between the electrons is, of course, always present, a pure Anderson transition cannot be expected to be realized in nature unless the interaction turns out to be irrelevant for the nature of the transition. As we shall see later, this in general is not the case. Unfortunately, this means that the comparatively simple models developed for disordered noninteracting electrons are insufficient for understanding experiments. On the other hand, most materials that display a metal-insulator transition are highly disordered, and the pure Mott transition picture is equally inadequate. Consider, for instance, the case of phosphorus-doped silicon, a particularly well studied example of a system showing a metal-insulator transition. Suppose one starts with pure silicon, which is an insulator at $T=0$. Upon doping with the donor phosphorus, extra electrons are brought into the system. However, for low dopant concentrations the overlap between donor states is exponentially small, and one expects an insulator with hydrogenlike impurity states. This is indeed what is observed (see Fig. 3). With increasing phosphorus concentration one finds a broadening of the lines due to impurity pairs, and finally a broad continuum due to a distribution of impurity cluster sizes, with the system still being an insulator (Fig. 3). With a further increase in phosphorus concentration, one

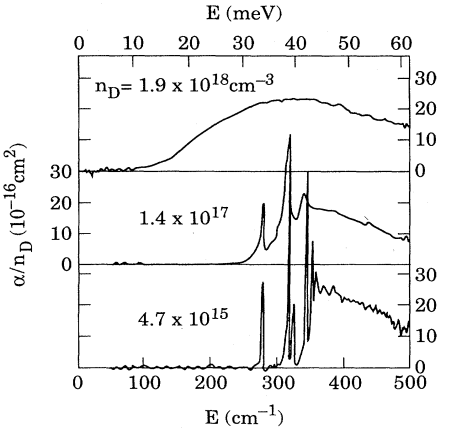

FIG. 3. Far-infrared absorption coefficient α for three different donor concentrations n_D in Si:P. The critical concentration in this system is $n_c \cong 3.7\times10^{18}$ cm^{-3}. From Thomas *et al.* (1981).

[1]The former behavior can be seen by choosing $b=t^{-\nu}$, the latter by choosing $b=\Omega^{-1/d}$. Scaling properties at the metal-insulator transition will be covered in detail in Sec. IV.

expects, according to Mott's argument, a transition to a metal at some critical concentration n_c. Again, this is what is observed (see Fig. 4). The observed transition is continuous and cannot be understood purely in terms of a Mott transition. The reason lies in the fact that the doping process not only introduces excess electrons, but at the same time creates disorder, since the dopant atoms are randomly distributed in the host lattice. One therefore would expect the transition, in part, to follow the Anderson model.

The conclusion that has been reached over the years is that neither Anderson's nor Mott's picture by itself is sufficient to understand the observed metal-insulator transition. Rather, one has to deal simultaneously with disorder and interactions between the electrons, neither one of which is a small effect near the transition. This has proven to be a very hard problem, which is far from having been solved completely. Somewhat ironically, the most precise experiment, viz., the one on Si:P shown in Fig. 4, has proven the hardest to understand for reasons that will be discussed in detail in Sec. VI. Nevertheless, substantial progress has been achieved in our understanding of this "Anderson-Mott transition."

An important development in this respect was the work of Finkel'stein (1983a, 1984a, 1984b), who extended the field-theoretic description of the Anderson transition (Wegner, 1979; Efetov *et al.*, 1980) to allow for interactions. This model not only allowed for the use of renormalization-group methods to deal with strong disorder, but also was able to consider interactions of arbitrary strength. It thus achieved two important improvements over previous perturbative work and quickly led to a description of the Anderson-Mott transition in the presence of magnetic impurities or a magnetic field (Finkel'stein, 1984a), which will be reviewed in Sec. V. These results were soon supplemented by a derivation in terms of resummed many-body perturbation theory

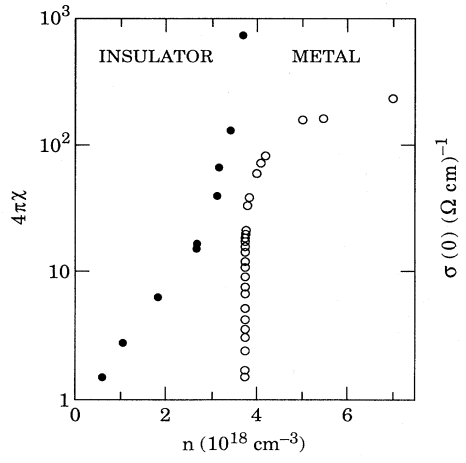

FIG. 4. Divergence of the dielectric susceptibility χ (●), and vanishing of the static conductivity σ (○), both extrapolated to zero temperature, at the metal-insulator transition in Si:P n is the P concentration. After Rosenbaum *et al.* (1983).

(Castellani, Di Castro, Lee, and Ma, 1984) and by interpretations in terms of Fermi-liquid theory (Altshuler and Aronov, 1983; Castellani and Di Castro, 1986).

In contrast to these successes, an understanding of the metal-insulator transition in the absence of either magnetic impurities or magnetic fields has proven much harder. The difficulties are twofold: the interaction amplitude in the particle-hole spin-triplet channel scales to infinity if it is not cut off by magnetic effects, and the particle-particle or Cooper interaction channel has a structure that is not easily amenable to standard renormalization-group techniques. The first problem has received much attention (Finkel'stein, 1983a, 1984b, 1984c; Castellani, Di Castro, Lee, Ma, Sorella, and Tabet, 1984, 1986; Castellani, Kotliar, and Lee, 1987) and was originally interpreted as being related to local moment formation, or as signaling an exotic metal-insulator transition in which the scaled disorder flows to zero at the transition. More recent work has suggested that it actually signals the presence of a phase transition that is magnetic in nature and distinct from the metal-insulator transition (Kirkpatrick and Belitz, 1990b, 1992b; Belitz and Kirkpatrick, 1991). The second problem has been considered (Castellani, Di Castro, Forgacs, and Sorella, 1984; Finkel'stein, 1984b; Kirkpatrick and Belitz, 1993), but the proposed solutions so far are not mutually consistent and cannot even tentatively be considered final. They will be discussed in Secs. V and VI. Another problem is that experiments on doped semiconductors (Paalanen *et al.*, 1986, 1988) show thermodynamic anomalies that cannot be consistently explained within the field-theoretic model and have prompted rather different theoretical approaches (Bhatt and Fisher, 1992). These will be considered in Sec. IX.

For all these reasons the metal-insulator-transition problem cannot be considered solved. However, the progress made within the last decade has not been reviewed, and it is the purpose of the present article to describe the current state of affairs. In doing so, one difficulty is that the problem has been tackled by a large variety of approaches that are very different with respect to both the underlying physical ideas and the technical methods used. On the physical side, one can distinguish between phenomenological approaches, on the one hand, which try to get clues from experiments about what physical effects are important near the transition and must be included in the theory, and what one might call the slow-mode philosophy on the other hand. The latter starts from the assumption that the physics near the metal-insulator transition will be dominated by the low-lying excitations of the system, which can be extracted from a simplified microscopic model. On the technical side, intuitive phenomenology, many-body perturbation theory, the renormalization group, and effective field-theoretic techniques have all played an important role. Since the problem remains unsolved, one cannot really afford the luxury of taking one of these points of view exclusively. We shall focus, however, on the low-lying-mode philoso-

phy, implemented by field-theoretic techniques, for two reasons. First, this line of approach is relatively new for interacting systems and has recently led to developments that have not been covered in previous reviews. Secondly, we believe that these techniques have the best chance of eventually providing us with a complete, microscopic theory of the metal-insulator transition. One should keep in mind, however, that even if one accepts the slow-mode approach, it is in general still an open problem how to determine all of the relevant slow modes. We shall come back to this problem in Secs. II, III, and X.

The plan of this review is as follows. We start in Sec. II with an elementary discussion of the slow modes, i.e., the diffusive modes that result from the conservation laws for particle number, spin, and energy. All of the material presented in that section can be found in various books and review articles. We feel, however, that our discussion is necessary both for pedagogical reasons and to put the results of the field theory presented later in the proper context. Section III is devoted to an explanation of the technical apparatus that will be used in most of the rest of the paper. That section is rather technical and extensive for two reasons: the field-theoretic methods underlying much of the work to be reviewed are not as widely known among condensed-matter theorists as, say, Green's-function techniques, and the details of the derivation of the fundamental model describing interacting disordered electrons have never been published. The section is written for readers who wish to work actively with the field theory. Anybody who is mainly interested in learning about the results presented in the later sections can skip over most of the technical details in Sec. III. Section IV is devoted to a general discussion of possible scaling scenarios for a metal-insulator transition of interacting electrons, i.e., the question of how to generalize Eq. (1.3) to the interacting case. Sections V and VI review explicit calculations that show how these scaling scenarios are realized in various universality classes. Section VII is devoted to a related subject, namely, the destruction of (conventional bulk) superconductivity near the metal-insulator transition. This problem is actually part of a more general one, namely, the question of how collective phenomena like superconductivity, magnetism, etc., are affected by strong disorder in the vicinity of a metal-insulator transition. Since the answer obviously requires a solution of the problem in the absence of the collective phenomenon, these issues have only recently started to be addressed. In Sec. VIII we discuss a recent suggestion of disorder-induced spin-triplet superconductivity in 2-d systems. In its existing form the slow-mode field theory is unlikely to accomplish the ultimate goal of providing a complete microscopic theory of all phenomena observed close to the metal-insulator transition. Rather, it will have to be supplemented by physical ideas developed through other approaches. Some of these are discussed in Sec. IX. Section X provides a summary and a discussion of what we consider to be the most pressing open problems in the field.

The localization problem has been reviewed previously a number of times, most notably by Altshuler and Aronov (1984), Bergmann (1984), Lee and Ramakrishnan (1985), Finkel'stein (1990), and MacKinnon and Kramer (1993). We have tried to avoid duplication of material as far as possible and often refer to these reviews rather than trying to be complete. With some exceptions, we also concentrate on the metal-insulator transition proper and its immediate vicinity, excluding effects at weak disorder or deep in the insulator. This holds in particular for our selection of experiments to be discussed in detail in Secs. V and VI. Throughout the paper we use units such that Planck's constant $\hbar$, Boltzmann's constant k_B, and minus the electron charge e, are equal to unity unless otherwise mentioned.

II. DIFFUSIVE ELECTRONS

A. Diffusion poles

The dynamics of conserved quantities show peculiarities that arise from the fact that, due to the conservation law, their values cannot change arbitrarily in space and time. Let us consider the density $n(\mathbf{x},t)$ of a conserved quantity N in some many-particle system. In equilibrium, n is constant in space and time: $n(\mathbf{x},t)\equiv n_0$. Suppose a fluctuation δn is created, $n(\mathbf{x},t)=n_0+\delta n(\mathbf{x},t)$, and we ask how the system will go back to equilibrium. Since n is conserved, it can do so only by transporting some N out of or into the region where $\delta n\neq 0$. If this region is large, this will take a long time. Therefore long-wavelength fluctuations of conserved quantities will decay very slowly. Slowly decaying fluctuations determine the low-lying modes and are of central importance for a description of the system. An example is classical fluid dynamics, where the conserved quantities are particle number, momentum, and energy, and the slow modes are first sound, heat diffusion, and transverse momentum diffusion (see, for example, Forster, 1975; Boon and Yip, 1980).

We shall be concerned with the dynamics of electrons moving in a random array of static scatterers. Since the scatterers can absorb momentum, the only conserved quantities are particle number (or charge), energy, and possibly spin. In general, all of these have diffusive dynamics. The central assumption of the theory we shall review is that the slow decay of charge, spin, and energy-density fluctuations leads to, and dominates, the physics near the metal-insulator transition. The basic strategy for a description at zero temperature is to start with perturbation theory in the diffusive phase and to study the instability of that phase. Of course, this presumes the existence of a diffusive phase somewhere in the phase diagram. If this is not true, one can use perturbation theory only at finite temperature, where transport is diffusive due to inelastic processes. As the temperature approaches zero, perturbation theory will then break

down everywhere in the parameter space. An example may be seen in the 2-d systems, for which all theoretical approaches now agree that at $T=0$ electrons are in general never diffusive (Abrahams *et al.*, 1979). More recently it has been suggested that the spin dynamics may never be diffusive, not even for $d>2$ (Bhatt and Fisher, 1992). The possible consequences of this are currently not quite clear. However, in the simplest possible scenario the absence of spin diffusion would merely change the universality class (see Sec. III.B.4.c) of the metal-insulator transition. We shall discuss this proposition further in Sec. IX. Here we proceed under the assumption that for $d>2$ at $T=0$ there is a small-disorder phase where charge, spin, and energy are diffusive.

In this subsection we consider noninteracting electrons in an environment of elastic, spin-independent scatterers. In this case the conservation laws for spin and energy do not add anything to particle number conservation, and the spin and heat diffusion coefficients are the same as the charge or number diffusion coefficient.[2] For the spin diffusion coefficient this is obvious, and for the heat diffusion it has been shown by Chester and Thellung (1961), Castellani, DiCastro, and Strinati (1987), and Strinati and Castellani (1987). We can therefore restrict ourselves to a discussion of particle number diffusion. The situation changes, of course, as soon as the electron-electron interaction is taken into account; see Sec. III.B.3.d.

1. The quasiclassical approximation for electron transport

The basic building block of the theory is the diffusive density response of the electrons in the quasiclassical approximation. We first discuss three derivations of the density response, in order of increasing technical sophistication. For the time being, we consider noninteracting electrons with a Hamiltonian

$$\hat{H}=\sum_{\mathbf{k}}[\mathbf{k}^2/2m-\mu]\hat{a}_{\mathbf{k}}^{+}\hat{a}_{\mathbf{k}}+\frac{1}{V}\sum_{\mathbf{q}}u(\mathbf{q})\hat{\rho}^{+}(\mathbf{q})\ . \tag{2.1a}$$

Here $\hat{a}_{\mathbf{k}}^{+}$ and $\hat{a}_{\mathbf{k}}$ are creation and annihilation operators for electrons in state $\mathbf{k}$, μ is the chemical potential, we assume free electrons with mass m, and V is the system volume. u is a random potential whose strength is given by

$$u_0=\frac{1}{2\pi N_F\tau}\equiv n_i\{|u(\mathbf{q})|^2\}_{\rm dis}\ , \tag{2.1b}$$

with N_F the density of states (DOS) per spin at the Fermi level, and τ the elastic mean free time in the Boltzmann approximation. n_i is the scatterer density, whose appearance results from performing the ensemble average $\{\cdots\}_{\rm dis}$ over the randomly situated scattering centers (Edwards, 1958). For simplicity we assume pointlike scatterers (i.e., s-wave scattering only). $\hat{\rho}(\mathbf{q})$ is the density operator,

$$\hat{\rho}(\mathbf{q})=\sum_{\mathbf{k}}\hat{a}_{\mathbf{k}-\mathbf{q}/2}^{+}\hat{a}_{\mathbf{k}+\mathbf{q}/2}\ . \tag{2.1c}$$

Since we are dealing with noninteracting electrons, spin results only in trivial factors of two and can be suppressed. We shall add the Coulomb interaction later.

a. A phenomenological argument for diffusion

Let us start with a very simple phenomenological argument for diffusive density dynamics (see, for example, Forster, 1975). Of course this approach does not depend on microscopic details and also holds for interacting electrons as well as for systems outside the quasiclassical regime. Consider a macroscopic number-density fluctuation $\delta n(\mathbf{x},t)$. Particle number conservation implies the continuity equation

$$\frac{\partial}{\partial t}\delta n(\mathbf{x},t)+\vec{\nabla}\cdot\mathbf{j}(\mathbf{x},t)=0\ , \tag{2.2}$$

with $\mathbf{j}$ the (macroscopic) number current density. It is plausible to assume that for a slowly varying density the current is proportional to the negative gradient of the density,

$$\mathbf{j}(\mathbf{x},t)=-D\vec{\nabla}\delta n(\mathbf{x},t)\ . \tag{2.3}$$

The positive coefficient D is called the diffusion constant. More precisely, $\mathbf{j}$ should be expressed in terms of a chemical potential gradient and an Onsager coefficient, which in turn can be expressed in terms of a density gradient and the diffusion coefficient (see, for example, DeGroot and Mazur, 1962). Combination of Eqs. (2.2) and (2.3) yields Fick's law,

$$\left[\frac{\partial}{\partial t}-D\Delta\right]\delta n(\mathbf{x},t)=0\ . \tag{2.4}$$

We Fourier transform and find the solution,

$$\delta n(\mathbf{q},t)=\delta n(\mathbf{q},0)e^{-Dq^2t},\quad t>0\ , \tag{2.5a}$$

which displays the slow decay of long-wavelength fluctuations mentioned above. We define a Laplace transform in time by

$$\delta n(\mathbf{q},z)\equiv\pm i\int dt\,\Theta(\pm t)e^{izt}\delta n(\mathbf{q},t)\ ,$$

$$\pm\ \text{for}\ \mathrm{Im}z\gtrless 0\ , \tag{2.5b}$$

with complex frequency z. This yields

$$\delta n(\mathbf{q},z)=\frac{-\delta n(\mathbf{q},t=0)}{z\pm iDq^2}\ . \tag{2.5c}$$

Here $\delta n(\mathbf{q},z)$ as a function of z has a branch cut at $\mathrm{Im}z=0$ and two Riemann sheets. The physical sheet is the one with no singularities. The analytic continuation

[2]The same is true for number and heat diffusion in the classical Lorentz model mentioned in Sec. I. Quantum mechanics does not change this.

to the other sheet from above and below the real axis has a pole at $z=-iDq^2$ and $z=iDq^2$, respectively. This pole is called a diffusion pole. Its relation to diffusive dynamics is obvious from Eqs. (2.5).

b. Linear response and the Boltzmann equation

We now turn to a microscopic derivation of Eqs. (2.5). Again, the formal part of this subsection is valid for general systems. Suppose the deviation from equilibrium, δn, is created by an external chemical potential $\mu_{\rm ext}(\mathbf{x},t)$. Then the Hamiltonian contains a term

$$\hat{H}_{\rm ext}(t)=-\int d\mathbf{x}\,\hat{\rho}(\mathbf{x})\mu_{\rm ext}(\mathbf{x},t)\ , \tag{2.6}$$

where $\hat{\rho}$ is the density operator, Eq. (2.1c). Linear-response theory (Fetter and Walecka, 1971; Forster, 1975) then tells us that the change in the expectation value of $\hat{\rho}$, to linear order in $\mu_{\rm ext}$, is given by

$$\begin{aligned}\delta n(\mathbf{x},t)&\equiv\langle\hat{\rho}(\mathbf{x},t)\rangle-\langle\hat{\rho}(\mathbf{x},t)\rangle_{\mu_{\rm ext}=0}\\&=i\int_{-\infty}^{t}dt'\int d\mathbf{x}'\chi_{\rho\rho}(\mathbf{x},\mathbf{x}';t,t')\mu_{\rm ext}(\mathbf{x}',t')\ ,\end{aligned} \tag{2.7a}$$

with the density susceptibility

$$\chi_{\rho\rho}(\mathbf{x},\mathbf{x}';t,t')=\langle[\hat{\rho}^{+}(\mathbf{x},t),\hat{\rho}(\mathbf{x}',t')]\rangle\ . \tag{2.7b}$$

Here $[\hat{a},\hat{b}]=\hat{a}\hat{b}-\hat{b}\hat{a}$ for any two operators $\hat{a}$, $\hat{b}$. The averaging is performed with the unperturbed Hamiltonian. If we include the ensemble average in the definition of the brackets in Eqs. (2.7), the system is translationally invariant in space and time, and we have

$$\delta n(\mathbf{q},t)=i\int_{-\infty}^{t}dt'\chi_{\rho\rho}(\mathbf{q},t-t')\mu_{\rm ext}(\mathbf{q},t')\ , \tag{2.8a}$$

with

$$\chi_{\rho\rho}(\mathbf{q},t)=\langle[\hat{\rho}^{+}(\mathbf{q},t),\hat{\rho}(\mathbf{q},t=0)]\rangle\ . \tag{2.8b}$$

A Laplace transformation according to Eq. (2.5b) yields the causal density susceptibility, which is equal to minus Zubarev's (1960) commutator correlation function,

$$\begin{aligned}\chi_{\rho\rho}(\mathbf{q},z)&=\pm i\int dt\,\Theta(\pm t)e^{izt}\chi_{\rho\rho}(\mathbf{q},t)\\&=-\langle\langle\hat{\rho}^{+}(\mathbf{q});\hat{\rho}(\mathbf{q})\rangle\rangle_z\ ,\end{aligned} \tag{2.9a}$$

with the notation

$$\langle\langle\hat{A};\hat{B}\rangle\rangle_z\equiv\mp i\int dt\,\Theta(\pm t)e^{izt}\langle[\hat{A}(t),\hat{B}]\rangle \tag{2.9b}$$

for any operators $\hat{A}$, $\hat{B}$. $\chi_{\rho\rho}(\mathbf{q},z)$ has the usual properties of causal functions (see, for example, Forster, 1975), which we list here without derivations. Causality allows for a spectral representation,

$$\chi_{\rho\rho}(\mathbf{q},z)=\int\frac{d\Omega}{\pi}\frac{\chi''_{\rho\rho}(\mathbf{q},\Omega)}{\Omega-z}\ , \tag{2.10a}$$

with a spectral function

$$\begin{aligned}\chi''_{\rho\rho}(\mathbf{q},\Omega)&=\frac{1}{2i}[\chi_{\rho\rho}(\mathbf{q},\Omega+i0)-\chi_{\rho\rho}(\mathbf{q},\Omega-i0)]\\&=\frac{1}{2}\int dt\,e^{i\Omega t}\chi_{\rho\rho}(\mathbf{q},t)\ .\end{aligned} \tag{2.10b}$$

The retarded and advanced susceptibilities are given by

$$\chi^{R,A}_{\rho\rho}(\mathbf{q},\Omega)=\chi_{\rho\rho}(\mathbf{q},\Omega\pm i0)=\chi'_{\rho\rho}(\mathbf{q},\Omega)\pm i\chi''_{\rho\rho}(\mathbf{q},\Omega)\ , \tag{2.11a}$$

where

$$\chi'_{\rho\rho}(\mathbf{q},\Omega)=\frac{1}{2}[\chi_{\rho\rho}(\mathbf{q},\Omega+i0)+\chi_{\rho\rho}(\mathbf{q},\Omega-i0)]\ . \tag{2.11b}$$

$\chi''_{\rho\rho}$ is positive semidefinite and determines the energy dissipation in the system. It is therefore also called the "dissipative part" of the susceptibility. It is related to the "reactive part" $\chi'_{\rho\rho}$ by means of a Kramers-Kronig relation,

$$\chi'_{\rho\rho}(\mathbf{q},\Omega)=\int\frac{d\Omega'}{\pi}\frac{\chi''_{\rho\rho}(\mathbf{q},\Omega')}{\Omega'-\Omega}\ , \tag{2.12a}$$

$$\chi''_{\rho\rho}(\mathbf{q},\Omega)=-\int\frac{d\Omega'}{\pi}\frac{\chi'_{\rho\rho}(\mathbf{q},\Omega')}{\Omega'-\Omega}\ . \tag{2.12b}$$

Finally, the dissipative part $\chi''_{\rho\rho}$ is related to the spontaneous density fluctuations in the system by the fluctuation-dissipation theorem (Callen and Welton, 1951). Corresponding relations hold for the general $A-B$ susceptibility, Eq. (2.9b), and generally for any causal function $f(z)$ instead of $\chi_{\rho\rho}(\mathbf{q},z)$.

The susceptibility determines the response of the system to external perturbations. The exact form of the response still depends on the nature of the perturbation (Kubo, 1957). Suppose the perturbation is suddenly switched on at $t=0$: $\mu_{\rm ext}(\mathbf{q},t)=\Theta(t)\mu_{\rm ext}(\mathbf{q})$. Then one finds from Eq. (2.8a)

$$\delta n(\mathbf{q},z)=-\frac{1}{z}\chi_{\rho\rho}(\mathbf{q},z)\mu_{\rm ext}(\mathbf{q})\ . \tag{2.13}$$

Now suppose that the perturbation is turned on adiabatically at $t=-\infty$ and switched off at $t=0$: $\mu_{\rm ext}(\mathbf{q},t)=\Theta(-t)e^{\varepsilon t}\mu_{\rm ext}(\mathbf{q})$ $(\varepsilon\to 0)$. Then one finds

$$\delta n(\mathbf{q},z)=\Phi^{R}_{\rho\rho}(\mathbf{q},z)\mu_{\rm ext}(\mathbf{q})\ . \tag{2.14a}$$

Here $\Phi^{R}_{\rho\rho}$ is Kubo's relaxation function,

$$\Phi^{R}_{\rho\rho}(\mathbf{q},z)=\frac{1}{z}[\chi_{\rho\rho}(\mathbf{q},z)-\chi^{0}_{\rho\rho}(\mathbf{q})]\ , \tag{2.14b}$$

with the static susceptibility

$$\chi^{0}_{\rho\rho}(\mathbf{q})=\chi_{\rho\rho}(\mathbf{q},z=i0)=\int\frac{d\Omega}{\pi}\chi''_{\rho\rho}(\mathbf{q},\Omega)/\Omega\ . \tag{2.14c}$$

Kubo has noted that the static susceptibility χ^0 in general is different from the isothermal susceptibility,

$$\chi^{T}_{\rho\rho}(\mathbf{q})=\frac{\partial}{\partial\mu_{\rm ext}(\mathbf{q})}\langle\rho(\mathbf{q})\rangle\equiv\frac{\partial n}{\partial\mu}(\mathbf{q})\ , \tag{2.15}$$

which enters the Kubo function

$$\Phi_{\rho\rho}(\mathbf{q},z)=\frac{1}{z}\left[\chi_{\rho\rho}(\mathbf{q},z)-\frac{\partial n}{\partial\mu}(\mathbf{q})\right] . \tag{2.16a}$$

At $q=0$, the isothermal density susceptibility is related to the compressibility $\kappa=-(\partial V/\partial p)_T/V$ by (Forster, 1975)

$$\partial n/\partial\mu=n^2\kappa , \tag{2.16b}$$

so $(\partial n/\partial\mu)(\mathbf{q})$ can be interpreted as the wave-vector-dependent compressibility. Kubo (1957) has shown that for ergodic variables[3] one has $\chi^0=\chi^T$. In this review we shall encounter no nontrivial nonergodic variables, and we shall not distinguish between χ^0 and χ^T. The distinction is crucial, however, for a description of, e.g., the insulating side of the metal-insulator transition (Götze, 1981), and it may be important for a solution of some of the open problems discussed in Sec. X. Let us also mention that for free electrons one has $\chi^0_{\rho\rho}(\mathbf{q})=(\partial n/\partial\mu)(\mathbf{q})=g(\mathbf{q})$, with g the static Lindhard function (Lindhard, 1954; Pines and Nozières, 1989). In the homogeneous limit, $g(q=0)=N_F$. That is, for free electrons the compressibility and the single-particle DOS are the same. This is *not* so for interacting electrons, as can be seen already at the level of Fermi-liquid theory, where the compressibility contains the Landau parameter F_0^s, while N_F does not [see, for example, Pines and Nozières, 1989, and Eqs. (3.126) and (3.127) below]. In the presence of disorder, the distinction between the two quantities is crucial (Lee, 1982).

In order to determine the general form of the density response, we use the equations of motion for the susceptibility (Zubarev, 1960),

$$\begin{aligned} z\langle\langle\,\hat{A};\hat{B}\,\rangle\rangle_z&=\langle[\,\hat{A},\hat{B}\,]\rangle-\langle\langle\,[\hat{H},\hat{A}\,];\hat{B}\,\rangle\rangle_z \\ &=\langle[\,\hat{A},\hat{B}\,]\rangle+\langle\langle\,\hat{A};[\hat{H},\hat{B}\,]\rangle\rangle_z . \end{aligned} \tag{2.17}$$

For the density operator, Eq. (2.1c), we have

$$[\hat{H},\hat{\rho}(\mathbf{q})]=\mathbf{q}\cdot\hat{\mathbf{j}}(\mathbf{q}) , \tag{2.18a}$$

with the current-density operator

$$\hat{\mathbf{j}}(\mathbf{q})=\frac{1}{m}\sum_{\mathbf{k}}\mathbf{k}\hat{a}^+_{\mathbf{k}-\mathbf{q}/2}\hat{a}_{\mathbf{k}+\mathbf{q}/2} . \tag{2.18b}$$

Equations (2.18) are the microscopic analog of Eq. (2.2). Notice that they are a consequence of particle number conservation and remain valid for interacting electrons. Applying Eqs. (2.17) twice, and using $\langle[\hat{\mathbf{j}}^+(\mathbf{q}),\hat{\rho}(\mathbf{q})]\rangle=\mathbf{q}n/m$, we find

$$\chi_{\rho\rho}(\mathbf{q},z)=\frac{q^2}{z}\Phi_L(\mathbf{q},z) , \tag{2.19a}$$

with the longitudinal current Kubo function

$$\Phi_L(\mathbf{q},z)=\frac{1}{z}\left[-(1/q^2)\langle\langle\,\mathbf{q}\cdot\hat{\mathbf{j}}^+(\mathbf{q});\mathbf{q}\cdot\hat{\mathbf{j}}(\mathbf{q})\,\rangle\rangle_z-\frac{n}{m}\right] . \tag{2.19b}$$

Let us consider Eqs. (2.19) in the limit $q\to 0$. Kubo (1957) has shown that $\Phi_L(q=0,z)$ determines the dynamical conductivity,

$$\sigma(z)=-i\Phi_L(q=0,z) . \tag{2.20}$$

For small frequencies, σ is further related to the diffusion constant by an Einstein relation (Kubo, 1957),

$$\lim_{\Omega\to 0}\sigma(\Omega\pm i0)=\pm\frac{\partial n}{\partial\mu}D . \tag{2.21}$$

It follows from Eq. (2.19a) that in the hydrodynamic limit $\chi_{\rho\rho}$ vanishes like q^2 as a result of particle number conservation,

$$\lim_{q\to 0}\chi_{\rho\rho}(\mathbf{q},z)=\frac{q^2}{z}iD\frac{\partial n}{\partial\mu}\mathrm{sgn}(\mathrm{Im}z)+o(q^2) . \tag{2.22a}$$

Here $o(q^2)$ denotes terms that vanish faster than q^2 as $q^2\to 0$. If we compare Eq. (2.22a) with the zero-frequency result

$$\lim_{\Omega\to 0}\chi_{\rho\rho}(\mathbf{q},\Omega+i0)=\frac{\partial n}{\partial\mu}(\mathbf{q}) , \tag{2.22b}$$

we see that the limits $q\to 0$ and $\Omega\to 0$ do not commute, so $\chi_{\rho\rho}(\mathbf{q},z)$ must be nonanalytic at $q=0$, $z=0$. Indeed, we can use Eq. (2.22a) to write the Kubo function in the hydrodynamic limit as

$$\Phi_{\rho\rho}(q\to 0,z)=\frac{-\partial n/\partial\mu}{z+iDq^2\mathrm{sgn}(\mathrm{Im}z)} . \tag{2.23}$$

We thus recover the diffusion pole of Eq. (2.5c). Note that the Kubo function is the appropriate response function to compare with Eq. (2.5c), since in our phenomenological argument we had assumed an adiabatically prepared nonequilibrium state, which was allowed to relax according to the unperturbed system's dynamics. For later reference we note that Eqs. (2.21)–(2.23) are generally valid, not just for noninteracting electrons.

The question remains how to calculate σ or D. In the quasiclassical approximation, we can use the Boltzmann equation with the result (see, for example, Ziman, 1964)

$$D=\frac{n/m}{\partial n/\partial\mu}\tau , \tag{2.24a}$$

where τ is the elastic mean free time. In three dimensions,

$$\frac{1}{\tau}=n_i 2\pi v_F\int_0^\pi d\vartheta\,\sin\vartheta(1-\cos\vartheta)\sigma(\vartheta) , \tag{2.24b}$$

where $\sigma(\vartheta)$ denotes the differential scattering cross sec-

[3] A nonergodic variable in this sense is one whose correlations do not decay in the limit of long times, so its Kubo function diverges for small z like $\Phi(q,z\to 0)=-f(q)/z$. Notice that this is always the case for conserved quantities at $q=0$, but is a nontrivial property at $q\neq 0$. For a mathematical discussion of ergodicity, see Khinchin (1949).

tion. For isotropic scattering the $\cos\vartheta$ term does not contribute, and if we treat the scattering process in the Born approximation we recover Eq. (2.1b). We can also solve the Boltzmann equation directly for the density response. In general one obtains Eq. (2.23) for the density response function. For the special case of isotropic scattering the explicit result for the diffusion coefficient is again given by Eqs. (2.24) (see, for example, Hauge, 1974).

c. Diagrammatic derivation of the diffusive density response

We shall now calculate the density and current correlation functions explicitly by means of many-body perturbation theory. In order to do so, we have to rewrite the commutator correlation function, Eq. (2.8b), in terms of a time-ordered correlation function. This can be done using standard techniques (Fetter and Walecka, 1971; Mahan, 1981) and allows for a convenient handling of the correlation function formalism at finite temperatures. We define an imaginary-time correlation function in the Matsubara formalism,

$$\pi_{\rho\rho}(\mathbf{q},\tau)=-\langle T_\tau\hat{\rho}^+(\mathbf{q},\tau)\hat{\rho}(\mathbf{q},\tau=0)\rangle\ , \tag{2.25a}$$

and its Fourier transform

$$\pi_{\rho\rho}(\mathbf{q},i\Omega_n)=\int_0^\beta d\tau\, e^{i\Omega_n\tau}\pi_{\rho\rho}(\mathbf{q},\tau)\ . \tag{2.25b}$$

Here τ denotes imaginary time, T_τ is the imaginary-time ordering operator, $\Omega_n=2\pi Tn$, with n an integer, is a bosonic Matsubara frequency, and $\beta=1/T$ is the inverse temperature. $\pi_{\rho\rho}(\mathbf{q},i\Omega_n)$, which is often called the polarization function, is identical to minus the causal density susceptibility, Eq. (2.10a), taken at $z=i\Omega_n$. The retarded and advanced susceptibilities, Eq. (2.11a), can be obtained by analytical continuation to real frequencies,

$$\pi^{R,A}(\mathbf{q},\Omega)=\pi(\mathbf{q},i\Omega_n\to\Omega\pm i0)\ . \tag{2.26}$$

An analogous polarization function can be formed with the current operator. The Kubo formula for the conductivity, Eq. (2.20), then takes the form

$$\sigma(\Omega)=i\frac{1}{i\Omega_n}\left[\pi_L(q=0,i\Omega_n)+\frac{n}{m}\right]\Bigg|_{i\Omega_n\to\Omega+i0}\ , \tag{2.27a}$$

where

$$\pi_L(\mathbf{q},i\Omega_n)=-\int_0^\beta d\tau\, e^{i\Omega_n\tau}\langle T_\tau(\mathbf{q}/q)\cdot\hat{\mathbf{j}}^+(\mathbf{q},\tau)(\mathbf{q}/q)\cdot\hat{\mathbf{j}}(\mathbf{q},\tau=0)\rangle\ . \tag{2.27b}$$

The Wick theorem can now be used to evaluate the time-ordered representations of the correlation functions in perturbation theory.

For our present purposes, the small parameter for a perturbative treatment is the density of scatterers n_i. The averaging over the random positions of the scatterers can be performed using the technique developed by Edwards (1958; see also Abrikosov *et al.*, 1975; Mahan, 1981). The building blocks of the theory are, first, the bare-electron Green's function,

$$G^{(0)}(\mathbf{q},i\omega_n)=-\int_0^\beta d\tau\, e^{i\omega_n\tau}\langle T_\tau\hat{a}_{\mathbf{q}}(\tau)\hat{a}_{\mathbf{q}}^+(\tau=0)\rangle_{\hat{H}_0}\ , \tag{2.28}$$

and, second, the impurity factor u_0, Eq. (2.1b). Here $\omega_n=2\pi T(n+1/2)$, with n an integer, is a fermionic Matsubara frequency, and the index $\hat{H}_0$ indicates that the average is to be taken with the free-electron part of the Hamiltonian only. Diagrammatically, we denote $G^{(0)}$ by a directed light line, and u_0 by two broken lines [one for each factor of $u(\mathbf{q})$] and a cross (for the factor of n_i); see Fig. 5(a). To zeroth order in the impurity density, the density polarization function is given by

$$\pi^{(0)}_{\rho\rho}(\mathbf{q},i\Omega_n)=\sum_{\mathbf{p}}T\sum_{i\omega_n}G^{(0)}(\mathbf{p},i\omega_n)\times G^{(0)}(\mathbf{p}+\mathbf{q},i\omega_n+i\Omega_n)\ , \tag{2.29}$$

which is shown graphically in Fig. 5(b). With the explicit expression for $G^{(0)}$,

$$G^{(0)}(\mathbf{p},i\omega_n)=[i\omega_n-\mathbf{p}^2/2m+\mu]^{-1}\ , \tag{2.28'}$$

the integrals are easily done, and one obtains the familiar Lindhard function (Lindhard, 1954). The same result can be obtained by evaluating the commutator in Eq. (2.7b) and performing the Fourier-Laplace transform. For finite impurity concentrations, we know from the previ-

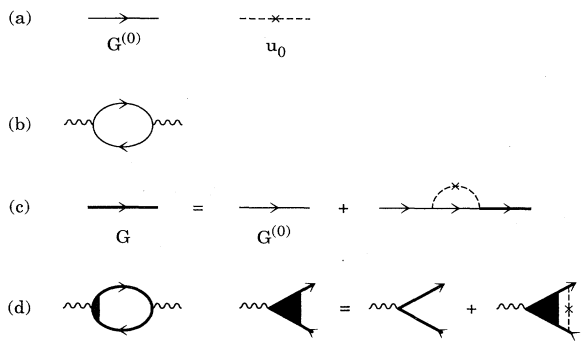

FIG. 5. Diagrammatic elements of perturbation theory: (a) diagrammatic representation of the bare Green's function and the impurity factor; (b) diagrammatic representation of the bare density polarization function; (c) the Green's function in the Born approximation; (d) conserving approximation for the density polarization function.

ous subsection that the diffusion constant goes like n_i^{-1} for small n_i. It is therefore clear that any expansion in powers of n_i will require an infinite resummation in order to reproduce the diffusion pole. In order not to violate particle number conservation, one actually has to do two separate infinite resummations. The first is to dress the bare Green's function by means of the Born approximation shown in Fig. 5(c). The result is

$$G(\mathbf{p},i\omega_n)=[i\omega_n-\mathbf{p}^2/2m+\mu+\Sigma(\mathbf{p},i\omega_n)]^{-1}\ , \tag{2.30a}$$

with the self-energy in the Born approximation

$$\Sigma(\mathbf{p},i\omega_n)=\frac{i}{2\tau}\mathrm{sgn}(\omega_n)\ . \tag{2.30b}$$

If one substituted Eqs. (2.30) for $G^{(0)}$ in Eq. (2.29), one would obtain a result that violates particle number conservation or gauge invariance. In fact, it is well known from quantum electrodynamics (Koba, 1951) that in order to maintain gauge invariance one has to treat vertex corrections consistently with self-energy corrections. This is the case in the approximation shown in Fig. 5(d), which reads

$$\pi_{\rho\rho}(\mathbf{q},i\Omega_n)=\sum_{\mathbf{p}}T\sum_{i\omega_n}\Gamma_\rho(\mathbf{p},\mathbf{q},i\omega_n,i\Omega_n)G(\mathbf{p}+\mathbf{q}/2,i\omega_n+i\Omega_n)G(\mathbf{p}-\mathbf{q}/2,i\omega_n)\ . \tag{2.31a}$$

The density vertex obeys the equation

$$\Gamma_\rho(\mathbf{p},\mathbf{q};i\omega_n,i\Omega_n)=1+u_0\sum_{\mathbf{k}}\Gamma_\rho(\mathbf{k},\mathbf{q};i\omega_n,i\Omega_n)G(\mathbf{k}+\mathbf{q}/2,i\omega_n+i\Omega_n)G(\mathbf{k}-\mathbf{q}/2,i\omega_n)\ , \tag{2.31b}$$

which is also shown graphically in Fig. 5(d). Notice that in a calculation of π_L rather than $\pi_{\rho\rho}$ (see, for example, Mahan, 1981), the bare vertex is given by $\mathbf{p}\cdot\mathbf{q}/q$ rather than by 1, and hence the vertex corrections vanish. Moreover, if we had not assumed pointlike scatterers the impurity factor u_0 would be momentum dependent and the integral equation (2.31b) would not be separable. With u_0 simply a number, Γ_ρ is independent of $\mathbf{p}$, and we have

$$\Gamma_\rho(\mathbf{q};i\omega_n,i\Omega_n)=(1-I_0(\mathbf{q};i\omega_n,i\Omega_n))^{-1}\ , \tag{2.32a}$$

where I_0 is the first in a set of integrals,

$$I_m(\mathbf{q};i\omega_n,i\Omega_n)=\frac{1}{2\pi N_F\tau}\sum_{\mathbf{k}}(\mathbf{k}\cdot\mathbf{q}/kq)^m G(\mathbf{k}+\mathbf{q}/2,i\omega_n+i\Omega_n)G(\mathbf{k}-\mathbf{q}/2,i\omega_n),\quad m=0,1,\ldots\ . \tag{2.32b}$$

Even with the simple approximation, Eqs. (2.30), for G the integrals cannot be obtained in closed form, but their behavior for $\tau\to\infty$ can be determined systematically (Kirkpatrick and Belitz, 1986a). To lowest order in $1/\tau$ the following simple replacement is sufficient (Abrikosov *et al.*, 1975):

$$\begin{aligned}I_m(\mathbf{q};i\omega_n,i\Omega_n)&\simeq\frac{1}{2\pi N_F\tau}N_F\int d\varepsilon_{\mathbf{k}}\frac{1}{2}\int_{-1}^{1}dx\frac{x^m}{i(\omega_n+\Omega_n)-\varepsilon_{\mathbf{k}}-(k_Fq/m)x+\frac{i}{2\tau}\mathrm{sgn}(\omega_n+\Omega_n)}\frac{1}{i\omega_n-\varepsilon_{\mathbf{k}}+\frac{i}{2\tau}\mathrm{sgn}(\omega_n)}\\&=\Theta[-\omega_n(\omega_n+\Omega_n)]\frac{1}{2}\int_{-1}^{1}dx\frac{x^m}{1+|\Omega_n|\tau+iqlx\ \mathrm{sgn}(\Omega_n)}\ .\end{aligned} \tag{2.33a}$$

Here $l=v_F\tau$ is the mean free path, and we have assumed a 3-d system. For $d=2$, only the angular integration is different. In the limit of small q and Ω_n we have

$$I_0(\mathbf{q};i\omega_n,i\Omega_n)=\Theta[-\omega_n(\omega_n+\Omega_n)](1-|\Omega_n|\tau-Dq^2\tau)+O(\Omega_n^2,q^4,\Omega_nq^2) \tag{2.33b}$$

and

$$I_2(\mathbf{q};i\omega_n,i\Omega_n)=\Theta[-\omega_n(\omega_n+\Omega_n)]\frac{1}{3}\left[1-|\Omega_n|\tau-\frac{3}{5}Dq^2\tau\right]+O(\Omega_n^2,q^4,\Omega_nq^2)\ , \tag{2.33c}$$

where $D=v_F^2\tau/d$ is the semiclassical diffusion constant, Eq. (2.24a), for free electrons. For the density polarization function in the hydrodynamic limit we then obtain

$$\pi_{\rho\rho}(\mathbf{q},i\Omega_n)=-\frac{\partial n}{\partial\mu}+\frac{\partial n}{\partial\mu}\frac{|\Omega_n|}{|\Omega_n|+Dq^2}\ . \tag{2.34}$$

The first term in Eq. (2.34) comes from the region in frequency space where $I_0=0$ and $\Gamma_\rho=1$. Apart from corrections of order $|\Omega_n|$ and $1/\tau$ this contribution is given by the static response of free electrons. The second term comes from the region $\omega_n(\omega_n+\Omega_n)<0$ and is therefore proportional to $|\Omega_n|$. For the density susceptibility

we finally obtain

$$\chi_{\rho\rho}(\mathbf{q},i\Omega_n)=-\pi_{\rho\rho}(\mathbf{q},i\Omega_n)=\frac{\partial n}{\partial\mu}\frac{Dq^2}{|\Omega_n|+Dq^2}\ , \tag{2.35}$$

and for the Kubo function, Eq. (2.16a),

$$\Phi_{\rho\rho}(\mathbf{q},i\Omega_n)=\frac{-\partial n/\partial\mu}{i\Omega_n+iDq^2\mathrm{sgn}(\Omega_n)}\ . \tag{2.36}$$

For the conductivity, Eqs. (2.20) and (2.19), we recover the Einstein relation,

$$\sigma(i\Omega_n)=\frac{\partial n}{\partial\mu}D\ \mathrm{sgn}(\Omega_n)+O(\Omega_n)\ . \tag{2.37}$$

These results are identical with those obtained in the previous subsection. In particular, we recover the diffusion pole structure of the density response. We conclude that the summation of ladder diagrams for the density vertex, Fig. 5(d), reproduces the results of the quasiclassical Boltzmann equation.

Of course, the result for the conductivity, Eq. (2.37), can also be obtained by evaluating Eqs. (2.27) directly (see, for example, Mahan, 1981). The corresponding diagrams are shown in Fig. 6. The integral equation for the vertex function Γ in Fig. 6 is easily solved,

$$\Gamma_{\mathbf{k},\mathbf{p},i\omega_n}(\mathbf{q},i\Omega_n)\equiv\Gamma_{i\omega_n}(\mathbf{q},i\Omega_n)$$
$$=u_0[1-I_0(\mathbf{q};i\omega_n,i\Omega_n)]^{-1}\ . \tag{2.38}$$

Notice that for pointlike scatterers the vertex corrections do not contribute to the conductivity. This is because the current vertex, shown as a triangle in Fig. 6, is odd under parity.

d. Beyond the quasiclassical approximation

One way to go beyond the quasiclassical approximation is to include additional classes of diagrams. A particularly well studied contribution to the conductivity is obtained by replacing the "diffusion ladder" vertex Γ in Fig. 6 by the "crossed-ladder" or "Cooperon" vertex Λ shown in Fig. 7 (Gorkov *et al.*, 1979; Abrahams *et al.*, 1980). The result for Λ is (Vollhardt and Wölfle, 1980)

$$\Lambda_{\mathbf{k},\mathbf{p},i\omega_n}(\mathbf{q},i\Omega_n)=\Gamma_{i\omega_n}(\mathbf{k}+\mathbf{p},i\Omega_n),\quad \mathbf{k}\simeq-\mathbf{p}\ . \tag{2.39}$$

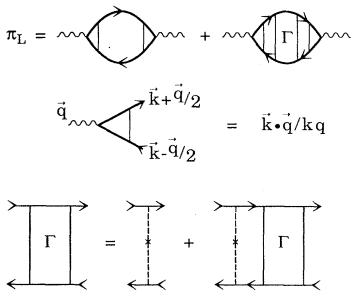

FIG. 6. Conserving approximation for the current polarization function.

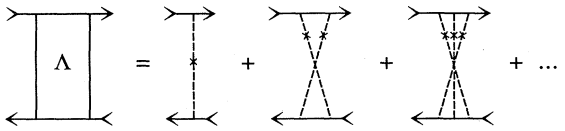

FIG. 7. Crossed-ladder vertex.

With Λ instead of Γ in Fig. 6 one obtains for the conductivity

$$\sigma(\Omega)=\sigma_0\left\{1-\frac{1}{\pi N_F}\int_{\mathbf{q}}\frac{1}{-i\Omega+Dq^2}\right\}\ , \tag{2.40}$$

where $\sigma_0=ne^2\tau/m$ is the quasiclassical result and $\int_{\mathbf{q}}\equiv\int d\mathbf{q}/(2\pi)^d$. The integral over the diffusion pole leads, at $\Omega=0$, to a logarithmic divergence in $d=2$. This corroborates the suggestion (Abrahams *et al.*, 1979) that for 2-*d* noninteracting electrons the static conductivity vanishes for all values of the disorder. This phenomenon is known as "weak localization" and has generated a substantial body of literature, which has been reviewed by Lee and Ramakrishnan (1985). In Sec. III.B.4.a we shall discuss another, symmetry-related, argument for the absence of diffusion in $d\leq 2$.

For $d>2$ at $\Omega=0$, Eq. (2.40) gives a correction to the static conductivity. For free electrons, the integral diverges in the ultraviolet and has to be cut off. It has been argued (Kawabata, 1981; Wölfle and Vollhardt, 1982) that this cutoff should be proportional to the inverse mean free path, $1/l=1/v_F\tau$. The argument given was that the diffusive form of the vertex function holds only in the hydrodynamic region $ql<1$. This leads, in $d=3$, to a correction to σ_0 in the limit $k_Fl\gg 1$ that reads

$$\sigma(\Omega=0)=\sigma_0\left\{1-\frac{\mathrm{const}}{(k_Fl)^2}+O((k_Fl)^{-3})\right\}\ . \tag{2.41}$$

This result, though very popular (see, for example, Mott, 1990, and references therein), is incorrect. The reason is, first, that in the limit $k_Fl\gg 1$ there are diagrammatic contributions to $\sigma(\Omega=0)$ that are not included in Λ, and, second, these contributions are not restricted to the hydrodynamic limit. The diagrams that contribute to the expansion of the static conductivity have been identified by Kirkpatrick and Dorfman (1983). In $d=3$, the result of the calculation is (Kirkpatrick and Belitz, 1986a)

$$\sigma(\Omega=0)=\sigma_0\left\{1-\frac{2\pi}{3}\frac{1}{k_Fl}+\frac{\pi^2-4}{8}\frac{1}{(k_Fl)^2}\ln\frac{1}{k_Fl}\right.$$
$$\left.+O((k_Fl)^{-2})\right\}\ . \tag{2.42}$$

The leading correction to the Boltzmann result is linear in $1/k_Fl$. The next-leading term is nonanalytic (Langer and Neal, 1966), as it is in classical systems (van Leeuwen and Weyland, 1967), and only the third-leading term will

be of $0[(k_F l)^{-2}]$. The leading correction in Eq. (2.42) has recently been quantitatively confirmed by experiment. Adams *et al.* (1992) reanalyzed data by various groups on the mobility of electrons in dense neutral gases. In these systems the electron density is so low that the electron-electron interaction is negligible, and the mean free path can be controlled by changing the gas density. Figure 8 shows data on electrons in H_2 and He together with the prediction of Eq. (2.42). The current experimental accuracy is not sufficient to check the theoretical prediction of a logarithmic correction to the term of $O((k_F l)^{-2})$. If this should become feasible, theory would also have to provide the constant coefficient of the term $\sim(k_F l)^{-2}$ in order to extract the logarithm from a constant background. Efforts to calculate this coefficient are under way (Wysokinski *et al.*, 1994).

Equation (2.40) also predicts that $\sigma(\Omega)$ is a nonanalytic function of frequency, the behavior at small frequency being

$$\sigma(\Omega\to 0)=\sigma(\Omega=0)\{1+\text{const}\times\Omega^{(d-2)/2}\} \quad (2.43a)$$

for $2<d<4$. In the time domain this corresponds to a behavior of the current-current correlation function, Eq. (2.27b), at long (real) times t,

$$\pi_L(q=0,t\to\infty)\propto t^{-d/2} . \quad (2.43b)$$

Such an algebraic decay of autocorrelation functions is known as a long-time tail and is characteristic of disordered systems. It was first found numerically for classical hard-sphere-model fluids (Alder and Wainwright, 1970) and explained theoretically in terms of correlated collision events (Dorfman and Cohen, 1970; Ernst *et al.*, 1970). The existence of these long-time tails came as a surprise, especially the fact that they exist for arbitrarily low density, the regime of validity of the Boltzmann equation, which predicts an exponential decay of the current autocorrelation function. The salient point is that the Boltzmann equation becomes exact at fixed time in the limit of low density, but not at fixed density, no matter how small, in the limit of long times. Physically the existence of power-law decays implies that there is no separation of time scales in general transport theory. In the classical hard-sphere fluid the long-time-tail exponent is $d/2$ as in Eq. (2.43b), and the same is true in more realistic classical fluids (Pomeau and Résibois, 1975; Forster *et al.*, 1977). In the classical Lorentz model mentioned in Sec. I, the exponent is $(d+2)/2$ (Ernst and Weyland, 1971). This difference, as well as a different sign of the prefactor of the long-time tail, is due to the missing dynamics of the scatterers in the Lorentz model. In a quantum Lorentz model, which is an appropriate model for localization of noninteracting electrons, the exponent has been shown to be $d/2$, and the prefactor has been calculated exactly (Kirkpatrick and Dorfman, 1983). The crossed-ladder approximation, Eq. (2.40), thus reproduces the exponent correctly. The physical reason for the different exponents in the classical and quantum Lorentz models is not entirely clear.

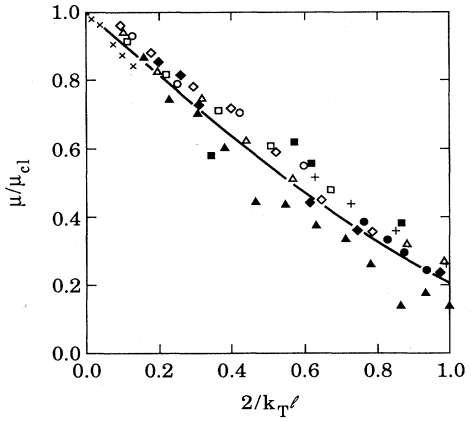

FIG. 8. Mobility μ of electrons in dense gases, normalized to the classical value μ_{cl}, as a function of the inverse mean free path. $k_T=\sqrt{2mT}$ is the thermal wave number. The symbols represent experimental values; the solid line is the theoretical result, Eq. (2.42). After Adams *et al.* (1992).

2. Disorder renormalization of electron-electron and electron-phonon interactions

Ultimately, we are interested in the interplay between the diffusion processes inherent in the vertex functions Γ and Λ and the Coulomb interaction between the electrons. For the theory of disordered superconductors (Sec. VII), as well as for the critical behavior of the sound attenuation described in Sec. V, we shall also need the diffusion corrections to the electron-phonon interaction. The theory of disordered superconductors is very complicated, since slow diffusive electron dynamics change the effective interaction, which in turn changes the electron dynamics. The sound attenuation is simpler because one can neglect the feedback of the phonons on the electron system. In this section we use simple diagrammatic perturbation theory to study the influence of the diffusion pole discussed in Sec. II.A.1 above on the dynamically screened Coulomb potential and on the electron-phonon interaction, neglecting all feedback effects.

a. Dynamical screening of diffusive electrons

In a many-electron system, the Coulomb potential

$$v_c(\mathbf{q})=(1-\delta_{q,0})\pi(2/q)^{d-1},\quad d=2,3 , \quad (2.44a)$$

is screened by the dielectric function. The screened potential is given by

$$V(\mathbf{q},i\Omega_n)=v_c(\mathbf{q})/\varepsilon(\mathbf{q},i\Omega_n) , \quad (2.44b)$$

and within the random-phase approximation (RPA) the dielectric function is given by the density response (see Pines and Nozières, 1989),

$$\varepsilon(\mathbf{q},i\Omega_n)=1+v_c(\mathbf{q})\chi_{\rho\rho}(\mathbf{q},i\Omega_n) . \quad (2.44c)$$

If we use the diffusive response, Eq. (2.35), in Eqs. (2.44), we obtain for the dynamically screened potential at small frequencies and wave numbers

$$V(\mathbf{q},i\Omega_n)=\begin{cases}\dfrac{4\pi}{q^2}\dfrac{|\Omega_n|+Dq^2}{D\kappa_3^2+|\Omega_n|+Dq^2}, & d=3,\\ \dfrac{2\pi}{q}\dfrac{|\Omega_n|+Dq^2}{D\kappa_2 q+|\Omega_n|+Dq^2}, & d=2,\end{cases} \tag{2.45a}$$

where

$$\kappa_d=(\pi 2^{d-1}\partial n/\partial\mu)^{1/(d-1)},\quad d=2,3\ , \tag{2.45b}$$

is the screening wave number. A characteristic feature of screening by diffusive electrons in the RPA is that the statically screened potential is given by the disorder-independent Thomas-Fermi expression, while at nonzero frequency the Coulomb singularity persists for wave numbers $q^2<|\Omega_n|/D$. With increasing disorder D decreases, and the phase-space region where the bare Coulomb singularity is present expands. This reflects the fact that the slow electrons have increasing difficulty screening fast charge fluctuations. From the preceding discussion it is clear that the general form of Eq. (2.45a) follows from particle number conservation and should be very generally valid. Surprisingly, a recent attempt to confirm the $|\Omega_n|/q^2$ singularity experimentally was unsuccessful (Bergmann and Wei, 1989). We also mention that Eq. (2.45a) is valid only in the limit of small frequencies. At large frequencies, the density response approaches that of free electrons, and one recovers the plasmon, weakly damped by disorder (Belitz and Das Sarma, 1986). For all dimensionalities $d>2$ the plasmon has a nonvanishing frequency at zero wave number. In the language of field theory, it is a massive mode. $d>2$ is also necessary in order to have a metal-insulator transition, and according to the soft-mode paradigm explained at the beginning of Sec. II.A the plasmon will be irrelevant for the critical behavior at the metal-insulator transition. In $d=2$ the bare plasmon is soft and overdamped by disorder in the region of small q (Giuliani and Quinn, 1984; Gold, 1984), but it still decays much faster than a diffusion mode and is therefore still irrelevant in the limit of small wavelengths and frequencies.

The diffusion pole that enters the dynamically screened Coulomb potential via the density response will appear at every electron-Coulomb vertex. The latter is denoted by a black triangle in Fig. 5(d) and given by Eq. (2.32a). We shall discuss the consequences of this in Sec. II.B below.

b. Coupling of phonons to diffusive electrons

The electromagnetic field couples to density fluctuations in the electron system. According to the simple Fröhlich model (see, for example, Abrikosov *et al.*, 1975) the same is true for the acoustic phonon field. One might therefore expect the electron-phonon vertex to be diffusion enhanced in the same way as the electron-Coulomb vertex. This conclusion is incorrect, as has been shown by Pippard (1955). Pippard's result has been confirmed and expanded on by many workers (Holstein, 1959; Tsuneto, 1961; Eisenriegler, 1973; Schmid, 1973; Grünewald and Scharnberg, 1974, 1975), but a large number of papers in the literature have overlooked this point and obtained a diffusion enhancement. A summary of the resulting confusion has been given by Belitz (1987b). The physical reason for the absence of any diffusion enhancement is that if the ionic lattice undergoes thermal motion, the electrons will follow almost coherently because of the system's tendency to maintain local charge neutrality. Because of this coherent motion, an impurity in the lattice will not lead to a strong effect in the electron-phonon coupling. Using a unitary transformation to a frame of reference that moves locally with the ions, Schmid (1973) showed that indeed the leading terms in the electron-phonon interaction vanish. The remaining effective interaction arises from a coupling between the lattice strain and the electronic stress tensor,

$$\hat{H}_{e-p}=i\sum_{\mathbf{k},\mathbf{q}}\sum_{b}\Gamma^{(0)}_{e-p}(\mathbf{k},\mathbf{q})\hat{a}^{+}_{\mathbf{k}-\mathbf{q}/2}\hat{a}_{\mathbf{k}+\mathbf{q}/2}\times\sqrt{\omega_b(\mathbf{q})/2}[\hat{B}_b(\mathbf{q})+\hat{B}_b^{+}(-\mathbf{q})]\ , \tag{2.46a}$$

with a bare electron-phonon vertex,

$$\Gamma^{(0)}_{e-p}(\mathbf{k},\mathbf{q})=\frac{1}{m\sqrt{\rho_{\rm ion}\omega_b(\mathbf{q})}}[\mathbf{k}\cdot\mathbf{q}][\mathbf{k}\cdot\mathbf{e}_b(\mathbf{q})]\ . \tag{2.46b}$$

Here $\hat{B}_b^{+}(\mathbf{q})$ and $\hat{B}_b(\mathbf{q})$ are phonon creation and annihilation operators with wave vector $\mathbf{q}$ and polarization index b ($b=L,T$ for longitudinal and transverse phonons, respectively). $\omega_b(\mathbf{q})$ is the bare-phonon dispersion relation ($\omega_b(|\mathbf{q}|\to 0)=c_b|\mathbf{q}|$ with sound velocity c_b), $\rho_{\rm ion}$ is the ionic mass density, and $\mathbf{e}_b$ is the phonon polarization vector. Equation (2.46b) replaces the $\mathbf{k}$-independent Fröhlich vertex. As in Fröhlich theory, the bare vertex has to be screened. This is done in the RPA, as in Sec. II.A.2.a above, and is shown diagrammatically in Fig. 9(a). The result can be expressed in terms of the integrals

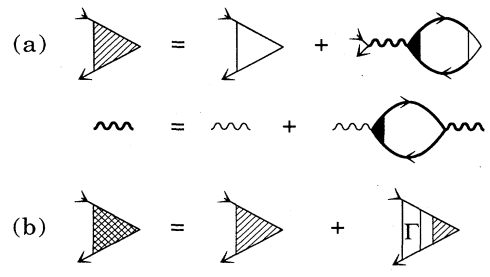

FIG. 9. The electron-phonon vertex. (a) The screened electron-phonon vertex (hatched triangle) in terms of the bare vertex (simple triangle) and the screened Coulomb potential (thick wavy line). The thin wavy line denotes the bare Coulomb potential, and the black triangle is defined in Fig. 5(d). (b) Impurity ladder corrections to the screened electron-phonon vertex.

I_m, Eq. (2.33a). In the longitudinal case one finds

$$\Gamma^{(s)}_{e-p}(\mathbf{k},\mathbf{q};i\omega_n,i\Omega_n)=\Gamma^{(0)}_{e-p}(\mathbf{k},\mathbf{q}) + \frac{k_F^2\delta_{b,L}}{3m\sqrt{\rho_{\rm ion}}c_b}\frac{1-I_0-3|\Omega_n|\tau I_2}{1-I_0-|\Omega_n|\tau I_0}\ , \tag{2.47}$$

where I_0 and I_2 are to be taken at arguments $\mathbf{q};i\omega_n,i\Omega_n$. There is no screening in the transverse case. The screened vertex has still to be corrected for impurities, Fig. 9(b), and one obtains the final result (Schmid, 1973)

$$\Gamma_{e-p}(\mathbf{k},\mathbf{q};i\omega_n,i\Omega_n) =\Gamma^{(0)}_{e-p}(\mathbf{k},\mathbf{p})+\frac{k_F^2\delta_{b,L}}{3m\sqrt{\rho_{\rm ion}}c_b}\left[\frac{1-I_0-3|\Omega_n|\tau I_2}{1-I_0-|\Omega_n|\tau I_0} -3I_2\right]\frac{1}{1-I_0}\ . \tag{2.48}$$

Explicit use of Eqs. (2.33) shows that, instead of a diffusion pole, the correction to the vertex has the structure $|\Omega_n|/(D\mathbf{q}^2+|\Omega_n|)$. This is finite if summed over $\mathbf{q}$, even in $d=2$ in the limit $|\Omega_n|\to 0$.

One can now use Eq. (2.48) to calculate the phonon self-energy $\pi_b(\mathbf{q},i\Omega_n)$, shown in Fig. 10, whose imaginary part determines the sound attenuation coefficient $\alpha_b(\mathbf{q})$ via

$$\alpha_b(\mathbf{q})=-\frac{\omega_b(\mathbf{q})}{c_b}{\rm Im}\pi_b[\mathbf{q},i\Omega_n\to\omega(\mathbf{q})+i0]\ . \tag{2.49}$$

In $d=3$ one recovers Pippard's (1955) result (Schmid, 1973),

$$\alpha_b(\mathbf{q})=[\omega_b^2(\mathbf{q})/c_b]d_b^2(3m^2/\pi^2)(1/\mathbf{q}^2l)f_b(|\mathbf{q}|l)\ , \tag{2.50a}$$

where $l=v_F\tau$ is the electronic mean free path, and

$$d_L=(c_T/c_L)d_T=k_F^2/3m\sqrt{\rho_{\rm ion}}c_L\ , \tag{2.50b}$$

and

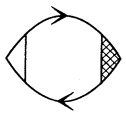

FIG. 10. The phonon self-energy.

$$f_L(x)=\frac{1}{3}\frac{x^2{\rm arctan}x}{x-{\rm arctan}x}-1\ , \tag{2.50c}$$

$$f_T(x)=\frac{1}{2x^3}[2x^3+3x-3(x^2+1){\rm arctan}x]\ . \tag{2.50d}$$

Notice that, in our simple jellium model, α_T vanishes in the clean limit, $|\mathbf{q}|l\gg 1$. Disorder enhances the coupling between electrons and transverse phonons (Keck and Schmid, 1976). In the long-wavelength or dirty limit, $|\mathbf{q}|l\ll 1$, one has

$$\alpha_L(\mathbf{q})=\alpha_T(\mathbf{q})(c_T/c_L)\frac{4}{3}=\frac{4}{45\pi^2}\frac{k_F^4}{c_L\rho_{\rm ion}}\mathbf{q}^2l,\quad d=3\ . \tag{2.51a}$$

In 2-d one obtains

$$\alpha_L(q)=\alpha_T(q)(c_T/c_L)=\frac{1}{8\pi^2}\frac{k_F^4}{c_L\rho_{\rm ion}}\mathbf{q}^2l,\quad d=2\ . \tag{2.51b}$$

In Eq. (2.51b), $\rho_{\rm ion}$ still denotes the bulk ion mass density.

An alternative route to these results has been given by Kadanoff and Falko (1964). They start out with the bare vertex, Eq. (2.46b), which shows that the sound attenuation is given by an electronic stress susceptibility $\chi_{\tau_b\tau_b}$, which is defined by Eqs. (2.9) with the density operator $\hat\rho$ replaced by the stress operator,

$$\hat\tau_b(\mathbf{q})=\sum_{\mathbf{k}}[\mathbf{k}\cdot\mathbf{q}/q][\mathbf{k}\cdot\mathbf{e}_b(\mathbf{q})]\hat a^+_{\mathbf{k}-\mathbf{q}/2}\hat a_{\mathbf{k}+\mathbf{q}/2}\ . \tag{2.52}$$

The stress susceptibility $\chi_{\tau_b\tau_b}$ still contains the electron-electron interaction, but not the electron-phonon interaction. The electron-electron interaction is then taken into account within the RPA as in Sec. II.A.2.a above. The result is

$$\alpha_b(\mathbf{q})=\frac{i\Omega_n}{m^2\rho_{\rm ion}c_b^3}{\rm Im}\{\chi_{\tau_b\tau_b}(\mathbf{q},i\Omega_n)-[\chi_{\tau_b\rho}(\mathbf{q},i\Omega_n)]^2/\chi_{\rho\rho}(\mathbf{q},i\Omega_n)\}|_{i\Omega_n\to\omega(\mathbf{q})+i0}\ , \tag{2.53}$$

where the correlation functions are now for noninteracting electrons. We see again that the screening corrections vanish in the transverse case, since there is no coupling between density and transverse stress. A calculation of the susceptibilities in Eq. (2.53) by means of the techniques mentioned in Secs. II.A.1.b, and II.A.1.c above leads again to Eqs. (2.50) and (2.51).

As in the case of the conductivity, one can go beyond the quasiclassical approximation by replacing the vertex function Γ in Fig. 9(b) by the crossed-ladder vertex Λ,

Fig. 7. These "weak-localization corrections" to the sound attenuation have been calculated by Houghton and Won (1985) and by Kotliar and Ramakrishnan (1985). The perturbative corrections to lowest order are similar to those to the conductivity, but an analysis of higher-order terms (Kirkpatrick and Belitz, 1986a) revealed that, unlike the case of the conductivity, the first-order perturbative result cannot simply be exponentiated to yield the critical behavior at the Anderson transition. The problem was solved by means of field-theoretic tech-

niques (Castellani and Kotliar, 1986, 1987, Kirkpatrick and Belitz, 1986b), which revealed the existence of two scaling parts for the sound attenuation. We shall come back to this in Sec. III.C.2.

B. Perturbation theory for interacting electrons; early scaling ideas

It is possible to study the interacting electron problem in perturbation theory by allowing for the dynamically screened Coulomb potential, Eq. (2.45a) (or simply a static, short-ranged model potential), together with the diffusion poles in the diagrammatic expansion. The simplest possibility is to account for the interaction within the RPA, neglecting the modification of the density vertex, Eq. (2.31b), by disorder. This approximation, which replaces the interacting problem by an effective noninteracting problem, has been studied by Gold and Götze (1983a, 1986) and has been used to analyze experiments (Gold and Götze, 1983b, 1985; Gold *et al.*, 1984; Gold 1985a, 1985b).

In order to see the qualitative modifications of the localization problem induced by the interaction, one must include diagrams that describe the interplay between interactions and disorder. To first order in both the interaction and the disorder this can still be done relatively easily. As an example, two contributions to the electronic self-energy are shown in Fig. 11. This interplay between diffusion and the Coulomb interaction was first considered by Schmid (1974), who found that the inelastic lifetime, i.e., the imaginary part of the self-energy, is enhanced by diffusion and shows a nonanalytic temperature or energy dependence. However, the subject became popular only after Altshuler and Aronov (1979a) found a corresponding nonanalyticity in the real part of the self-energy, which determines the tunneling density of states. In $d=2$, the nonanalyticity takes the form of a logarithm, and a logarithmic contribution was also found in the first-order correction to the conductivity (Altshuler, Aronov, and Lee, 1980). A remarkable aspect of this result was that, in the presence of interactions, ordinary diffusion ladders, Eq. (2.38), lead to the same kind of logarithmic divergence as do the crossed ladders, Eq. (2.39), in the case of noninteracting electrons. A large number of perturbative calculations followed, which further demonstrated the intimate interplay between disorder and interactions and laid the groundwork for later theories of the metal-insulator transition. This work has been covered in many reviews, e.g., Altshuler, Aronov, Khmelnitskii, and Larkin (1982), Altshuler and Aronov (1984), or Lee and Ramakrishnan (1985). There is no need to repeat this coverage here. To deduce from perturbation theory a scaling theory of the metal-insulator transition of interacting electrons proved very difficult, mainly due to the proliferation of the number of diagrams in the many-body formalism once one goes beyond first order in either the interaction or the disorder. This caused early attempts at this task to fail. However, after Finkel'stein (1983a) mapped the problem onto an effective field theory, the problem was also successfully dealt with within the many-body formalism, using the structure of the field theory as a guideline (Castellani, DiCastro, Lee, and Ma, 1984). Subsequently, the diagrammatic many-body approach was pursued in parallel with the field theory for a number of years, and it has played an important role in the physical interpretation of the coupling constants that appear in the field theory. We shall describe the results of this work together with those of the field-theoretic approach in the following sections.

FIG. 11. Two Coulomb contributions to the electron self-energy.

Before we turn to the field-theoretic description of the problem, let us mention the earliest attempt to construct a scaling theory, that of McMillan (1981). He used very general arguments to derive scaling equations that extended the one-parameter scaling of Abrahams *et al.* (1979) by adding an interaction strength as a new parameter. The metal-insulator transition was proposed to correspond to a fixed point at finite values of both disorder and interaction strength. It was pointed out by Lee (1982) that the paper contained a mistake [the Einstein relation, Eq. (2.37), was used with the single-particle DOS instead of $\partial n/\partial\mu$] which rendered invalid the relations between exponents derived by McMillan. However, his general scaling picture has proven to hold for all universality classes in which the interaction strength does not show runaway flow (Castellani, DiCastro, Lee, and Ma, 1984; Finkel'stein, 1984a).

III. FIELD-THEORETIC DESCRIPTION

A. Field theories for fermions

At the heart of the field-theoretic formulation of the quantum-mechanical many-body problem is an expression of the partition function in terms of a functional integral of the form

$$Z=\int D[\bar{\psi},\psi]e^{S[\bar{\psi},\psi]}\ . \tag{3.1}$$

Here $\psi=\psi(\mathbf{x},\tau)$ is an auxiliary field that depends on position $\mathbf{x}$ and imaginary time τ, and $\bar{\psi}$ is the conjugate field. For many-boson systems, ψ is complex valued and $\bar{\psi}$ is its complex conjugate. For fermion systems, we shall see that we need anticommuting (Grassmann) fields. S is referred to as the action, and we shall derive its functional form below.

A very efficient way to derive Eq. (3.1) is to use so-called coherent states, i.e., eigenstates of annihilation operators. [Coherent states were first introduced in

quantum electrodynamics (Glauber, 1963), hence the name.] Given the coherent states, the procedure for bosons (Casher *et al.*, 1968), is straightforward. For fermions, in trying to follow the same arguments, one immediately encounters the following problem. Let $\hat{a}_\alpha^+$ and $\hat{a}_\alpha$ be creation and annihilation operators for fermions with quantum number (or a set of quantum numbers) α. The $\hat{a}^+$ and $\hat{a}$ obey the anticommutation relations

$$[\hat{a}_\alpha,\hat{a}_\beta]_+=[\hat{a}_\alpha^+,\hat{a}_\beta^+]_+=0, \quad \forall\alpha,\beta , \tag{3.2a}$$

$$[\hat{a}_\alpha,\hat{a}_\beta^+]_+=\delta_{\alpha\beta} , \tag{3.2b}$$

with $[\hat{a},\hat{b}]_+=\hat{a}\hat{b}+\hat{b}\hat{a}$ for any operators $\hat{a}$, $\hat{b}$. If the algebra of eigenvalues is correctly chosen (see below), it follows from Eq. (3.2a) that one can find a common eigenvector for all $\hat{a}_\alpha$. Let that common eigenvector be $|\psi\rangle$. The eigenvalue equation is

$$\hat{a}_\alpha|\psi\rangle=\psi_\alpha|\psi\rangle . \tag{3.3}$$

Operating on both sides of Eq. (3.3) with $\hat{a}_\beta$, and using Eq. (3.2a), we find that the eigenvalues ψ_α are not c numbers, but rather they anticommute. We are therefore forced to consider anticommuting variables, often referred to as Grassmann numbers. In the following subsection we list as many (or rather, as few) properties of Grassmann variables as is necessary for our purposes. This coverage is not meant in any sense to be complete or mathematically precise. For either purpose the reader is referred to the books by Berezin (1966, 1987). Our exposition follows Negele and Orland (1988), Itzykson and Drouffe (1989), and Zinn-Justin (1989).

1. Grassmann variables

Let us consider a set of M objects ψ_α ($\alpha=1,\ldots,M$). (We shall assume $M<\infty$. Additional complications occur for infinite M; see Berezin, 1966.) Let there be an additive operation between the ψ, $\psi_\alpha+\psi_\beta$, and a multiplication with c numbers such that the distributive law holds. Let there further be an (associative) multiplication operation $\psi_\alpha\psi_\beta$ such that

$$\psi_\alpha\psi_\beta+\psi_\beta\psi_\alpha=0, \quad \forall\alpha,\beta . \tag{3.4}$$

Now consider the set $\mathbb{G}$ of all linear combinations of mononomials,

$$\begin{aligned} f&=f_0+\sum_\alpha f_1(\alpha)\psi_\alpha+\sum_{\alpha<\beta} f_2(\alpha,\beta)\psi_\alpha\psi_\beta \\ &\quad+\sum_{\alpha<\beta<\gamma} f_3(\alpha,\beta,\gamma)\psi_\alpha\psi_\beta\psi_\gamma+\cdots \\ &=\sum_{n=0}^{M}\frac{1}{n!}\sum_{\alpha_1,\ldots,\alpha_n} f_n(\alpha_1,\ldots,\alpha_n)\psi_{\alpha_1}\cdots\psi_{\alpha_n} , \end{aligned} \tag{3.5}$$

where the f_n are c-number valued, totally antisymmetric functions of their arguments. It is tedious but straightforward to prove the following statements: (1) The expansion of f given in Eq. (3.5) is unique. (2) $\mathbb{G}$ forms a group with respect to addition, and a vector space of dimension 2^M over $\mathbb{C}$. (3) Equation (3.4) permits the definition of a multiplication operation on $\mathbb{G}$ by

$$\begin{aligned} fg&=f_0g_0+\sum_\alpha[f_0g_1(\alpha)+f_1(\alpha)g_0]\psi_\alpha \\ &\quad+\tfrac{1}{2}\sum_{\alpha,\beta}[f_0g_2(\alpha,\beta)+f_1(\alpha)g_1(\beta)-f_1(\beta)g_1(\alpha) \\ &\quad+f_2(\alpha,\beta)g_0]\psi_\alpha\psi_\beta+\cdots . \end{aligned} \tag{3.6}$$

$\mathbb{G}$ is closed under multiplication, and fulfills all axioms for a graded algebra over $\mathbb{C}$. $\mathbb{G}$ is called the Grassmann algebra with M generators $\psi_1,\ldots,\psi_M$. It is also called the exterior algebra of the M-dimensional vector space over $\mathbb{C}$ generated by the ψ_α. (For further information see, for example, Chap. IX.9 of Nickerson *et al.*, 1959 or Chap. 13 of van der Waerden, 1970). On an algebra with $M=2m$, $m\in\mathbb{N}$, generators one can define an involution: choose m generators ψ_α and with each ψ_α associate an adjoint $\bar{\psi}_\alpha$ such that

$$\bar{\bar{\psi}}_\alpha=\psi_\alpha , \tag{3.7a}$$

$$\overline{a\psi_\alpha}=a^*\bar{\psi}_\alpha, \quad \forall a\in\mathbb{C} , \tag{3.7b}$$

where a^* denotes the complex conjugate of a,

$$\overline{\psi_\alpha+\psi_\beta}=\bar{\psi}_\alpha+\bar{\psi}_\beta \tag{3.7c}$$

and

$$\overline{\psi_\alpha\psi_\beta}=\bar{\psi}_\beta\bar{\psi}_\alpha . \tag{3.7d}$$

On $\mathbb{G}$ one can define differentiation and integration (Berezin, 1966). Left and right derivatives are defined by their actions on mononomials,

$$\begin{aligned} \frac{\vec{\partial}}{\partial\psi_\alpha}\psi_{\beta_1}\cdots\psi_{\beta_s}&=\delta_{\alpha\beta_1}\psi_{\beta_2}\cdots\psi_{\beta_s}-\delta_{\alpha\beta_2}\psi_{\beta_1}\psi_{\beta_3}\cdots\psi_{\beta_s} \\ &\quad+\cdots+(-)^{s-1}\delta_{\alpha\beta_s}\psi_{\beta_1}\cdots\psi_{\beta_{s-1}} , \end{aligned} \tag{3.8a}$$

$$\begin{aligned} \psi_{\beta_1}\cdots\psi_{\beta_s}\frac{\overleftarrow{\partial}}{\partial\psi_\alpha}&=\delta_{\alpha\beta_s}\psi_{\beta_1}\cdots\psi_{\beta_{s-1}} \\ &\quad-\delta_{\alpha\beta_{s-1}}\psi_{\beta_1}\cdots\psi_{\beta_{s-2}}\psi_{\beta_s}+\cdots \\ &\quad+(-)^{s-1}\delta_{\alpha\beta_1}\psi_{\beta_2}\cdots\psi_{\beta_s} . \end{aligned} \tag{3.8b}$$

The chain rule holds in its usual form. For integration, one defines a measure $d\psi_\alpha$ which satisfies

$$[d\psi_\alpha,d\psi_\beta]_+=[d\psi_\alpha,\psi_\beta]_+=0, \quad \forall\alpha,\beta , \tag{3.9a}$$

and defines definite integrals

$$\int d\psi_\alpha=0 , \tag{3.9b}$$

$$\int d\psi_\alpha\psi_\alpha=1 . \tag{3.9c}$$

Notice that, as with ordinary functions that vanish at infinity, the integral of a total differential (viz., $1=d\psi_\alpha/d\psi_\alpha$) vanishes. Also notice that, according to

this definition, integration is the same as differentiation. We give four integration formulas without proof (the proofs can be found in Berezin, 1966, Negele and Orland, 1988, or Zinn-Justin, 1989).

Proposition 1 (integration by parts): Let $f(\psi)$, $g(\psi)\in\mathbb{G}$. Then

$$\int d\psi f(\psi)\left[\frac{\vec{\partial}}{\partial\psi_\beta}g(\psi)\right]=\int d\psi\left[f(\psi)\frac{\overleftarrow{\partial}}{\partial\psi_\beta}\right]g(\psi)\ , \tag{3.10}$$

where $\psi=\{\psi_\alpha\}$, and $d\psi=d\psi_{\alpha_M}\cdots d\psi_{\alpha_1}$.

Proposition 2 (change of variables): Consider the linear substitution

$$\psi_\alpha=\sum_\beta m_{\alpha\beta}\eta_\beta\ , \tag{3.11a}$$

with $m_{\alpha\beta}\in\mathbb{C}$. Then

$$\int d\psi f(\psi)=(\det m^{-1})\int d\eta f(\psi(\eta))\ . \tag{3.11b}$$

Notice that the inverse of the Jacobian appears on the right-hand side.

Proposition 3 (Dirac delta function): The Grassmannian delta function defined by

$$\delta(\psi_\alpha,\psi_\beta)=\int d\psi_\gamma e^{-\psi_\gamma(\psi_\alpha-\psi_\beta)},\quad \alpha\neq\beta\neq\gamma\neq\alpha\ , \tag{3.12a}$$

has the property

$$\int d\psi_\beta\delta(\psi_\alpha,\psi_\beta)f(\psi_\beta)=f(\psi_\alpha)\ . \tag{3.12b}$$

Proposition 4 (Gaussian integral): Let m be a complex skew-symmetric matrix: $m_{\alpha\beta}=-m_{\beta\alpha}$. Then

$$\int d\psi\exp\left[\sum_{\alpha,\beta}\psi_\alpha m_{\alpha\beta}\psi_\beta\right]=[\det(2m)]^{1/2}\ . \tag{3.13a}$$

Corollary (generalized Gaussian integral): Consider a Grassmann algebra $\mathbb{G}$ with involution, with generators ψ_1, $\bar{\psi}_1$. Let η_i, $\bar{\eta}_i$ be the generators of a Grassmann algebra $\mathbb{G}'$, which is isomorphic to $\mathbb{G}$. Then for any invertible, complex matrix m,

$$\int\prod_\alpha d\psi_\alpha d\bar{\psi}_\alpha\exp\left[\sum_{\alpha,\beta}\bar{\psi}_\alpha m_{\alpha\beta}\psi_\beta+\sum_\alpha(\bar{\eta}_\alpha\psi_\alpha+\bar{\psi}_\alpha\eta_\alpha)\right]$$
$$=(\det m)\exp\left[-\sum_{\alpha,\beta}\bar{\eta}_\alpha(m^{-1})_{\alpha\beta}\eta_\beta\right]\ , \tag{3.13b}$$

where the integrand on the left-hand side is an element of the direct sum of $\mathbb{G}$ and $\mathbb{G}'$.

2. Fermion coherent states

We now return to our fermion system with annihilation and creation operators $\hat{a}_\alpha$ and $\hat{a}_\alpha^+$, respectively. We define a Grassmann algebra by associating a generator ψ_α with each $\hat{a}_\alpha$, and a conjugate generator $\bar{\psi}_\alpha$ with each $\hat{a}_\alpha^+$. We require

$$[\psi_\alpha,\hat{a}_\beta]_+=[\psi_\alpha,\hat{a}_\beta^+]_+=[\bar{\psi}_\alpha,\hat{a}_\beta]_+=[\bar{\psi}_\alpha,\hat{a}_\beta^+]_+=0\ ,\quad\forall\alpha,\beta\ , \tag{3.14a}$$

and

$$(\psi_\alpha\hat{a}_\beta)^+=\hat{a}_\beta^+\bar{\psi}_\alpha,\quad\forall\alpha,\beta\ . \tag{3.14b}$$

Then a common eigenvector of the $\hat{a}_\alpha$ can be written (Negele and Orland, 1988)

$$|\psi\rangle=\left[\exp\left[-\sum_\alpha\psi_\alpha\hat{a}_\alpha^+\right]\right]|0\rangle=\prod_\alpha(1-\psi_\alpha\hat{a}_\alpha^+)|0\rangle\ , \tag{3.15a}$$

where $|0\rangle$ denotes the ground state of the fermion system. It is not hard to see that $|\psi\rangle$ has the desired property, viz.,

$$\hat{a}_\alpha|\psi\rangle=\psi_\alpha|\psi\rangle\ . \tag{3.15b}$$

We list four properties of the coherent state; the proofs can be found in Negele and Orland (1988):

Proposition 1 (adjoint of coherent state):

$$\langle\psi|=\langle 0|\exp\left[-\sum_\alpha\hat{a}_\alpha\bar{\psi}_\alpha\right]\ , \tag{3.16a}$$

with

$$\langle\psi|\hat{a}_\alpha^+=\langle\psi|\bar{\psi}_\alpha\ . \tag{3.16b}$$

Proposition 2 (overlap of coherent states):

$$\langle\psi|\psi'\rangle=\exp\left[\sum_\alpha\bar{\psi}_\alpha\psi'_\alpha\right]\ . \tag{3.17}$$

Proposition 3 (completeness relation):

$$1=\int\prod_\alpha d\bar{\psi}_\alpha d\psi_\alpha e^{-\sum_\alpha\bar{\psi}_\alpha\psi_\alpha}|\psi\rangle\langle\psi|\ . \tag{3.18}$$

Proposition 4 (traces of operators): Let $\hat{A}$ be an operator in the Fock space spanned by the complete set $\{|n\rangle\}$. Then the trace of $\hat{A}$ can be written

$$\mathrm{tr}\,\hat{A}=\sum_n\langle n|\hat{A}|n\rangle$$
$$=\int\prod_\alpha d\bar{\psi}_\alpha d\psi_\alpha e^{-\sum_\alpha\bar{\psi}_\alpha\psi_\alpha}\langle-\psi|\hat{A}|\psi\rangle\ . \tag{3.19}$$

Here $\langle-\psi|$ denotes the state given by Eq. (3.16a) with $\bar{\psi}$ replaced by $-\bar{\psi}$.

Equation (3.18), where 1 on the left-hand side denotes the unit operator in Fock space, implies that the coherent states form a complete set in Fock space (they are actually overcomplete, since they span the enlarged Fock space defined over $\mathbb{G}$ rather than $\mathbb{C}$). Equation (3.19) is crucial for writing the partition function in the form of Eq. (3.1).

3. The partition function for many-fermion systems

With the help of the coherent states, Eq. (3.1) can now be derived as follows (Casher *et al.*, 1969; Negele and Or-

land, 1988). Consider a many-fermion system with Hamiltonian operator $\hat{H}=\hat{H}(\hat{a}_\alpha^+,\hat{a}_\alpha)$. We are interested in

$$Z=\mathrm{tr}\, e^{-\beta\hat{K}}=\int d\mu(\psi)e^{-\sum_\alpha \bar{\psi}_\alpha\psi_\alpha}\langle -\psi|e^{-\beta\hat{K}}|\psi\rangle \ . \tag{3.20}$$

Here $\hat{K}=\hat{H}-\mu\hat{N}$ with $\hat{N}$ the particle number operator. We have defined a measure $d\mu(\psi)=\prod_\alpha d\bar{\psi}_\alpha d\psi_\alpha$, and we have used the coherent-state representation, Eq. (3.19), for the trace. The strategy is now to discretize imaginary time by dividing the interval $[0,\beta]$ into N steps of size $\varepsilon=\beta/N$. At each step we insert a unit operator expressed in terms of coherent states, Eq. (3.18). This yields

$$Z=\lim_{N\to\infty}\int\left[\prod_{k=1}^N d\mu(\psi^k)\right]e^{-\sum_{k=1}^N\sum_\alpha\bar{\psi}_\alpha^k\psi_\alpha^k}\times\prod_{k=1}^N\langle\psi^{k-1}|e^{-\varepsilon\hat{K}(\hat{a}^+_\alpha,\hat{a}_\alpha)}|\psi^k\rangle \tag{3.21}$$

with $\langle\psi^0|=\langle-\psi^N|$. In the limit $N\to\infty$ it is sufficient to keep terms to first order in ε. Then we can expand and reexponentiate $\exp(-\varepsilon\hat{K})$ in Eq. (3.21). With the help of Eqs. (3.15b) and (3.16b) we find

$$Z=\lim_{N\to\infty}\int\left[\prod_{k=1}^N d\mu(\psi^k)\right]e^{S[\bar{\psi},\psi]}\ , \tag{3.22a}$$

where

$$S[\bar{\psi},\psi]=-\varepsilon\sum_{k=1}^N\left\{\sum_\alpha\bar{\psi}_\alpha^k\left[\frac{\psi_\alpha^k-\psi_\alpha^{k-1}}{\varepsilon}-\mu\psi_\alpha^{k-1}\right]+H(\bar{\psi}_\alpha^k,\psi_\alpha^{k-1})\right\}\ , \tag{3.22b}$$

with $\psi_\alpha^0=-\psi_\alpha^N$. Here $H(\bar{\psi}_\alpha^k,\psi_\alpha^{k-1})$ is the Hamiltonian $\hat{H}(\hat{a}_\alpha^+,\hat{a}_\alpha)$ with $\hat{a}_\alpha^+$ and $\hat{a}_\alpha$ replaced by $\bar{\psi}_\alpha^k$ and ψ_α^{k-1}, respectively.

Equations (3.22) represent the final result, but for notational convenience it is customary to introduce a continuum notation. In the limit $N\to\infty$ the set $\{\psi_\alpha^0,\ldots,\psi_\alpha^N\}$ defines a Grassmann-valued function $\psi_\alpha(\tau)$, $\tau\in[0,\beta]$, and it is natural to write

$$\bar{\psi}_\alpha^k\frac{1}{\varepsilon}[\psi_\alpha^k-\psi_\alpha^{k-1}]\equiv\bar{\psi}_\alpha(\tau)\frac{\partial}{\partial\tau}\psi_\alpha(\tau)\ , \tag{3.23a}$$

$$H(\bar{\psi}_\alpha^k,\psi_\alpha^{k-1})\equiv H[\bar{\psi}_\alpha(\tau),\psi_\alpha(\tau)]\ . \tag{3.23b}$$

With a functional integration measure defined by

$$D[\bar{\psi},\psi]\equiv\lim_{N\to\infty}\prod_{k=1}^N\prod_\alpha d\bar{\psi}_\alpha^k d\psi_\alpha^k\ , \tag{3.23c}$$

we can write

$$Z=\int D[\bar{\psi},\psi]e^{S[\bar{\psi},\psi]}\ , \tag{3.24a}$$

with

$$S[\bar{\psi},\psi]=\int_0^\beta d\tau\sum_\alpha\bar{\psi}_\alpha(\tau)\left[\frac{-\partial}{\partial\tau}+\mu\right]\psi_\alpha(\tau)-\int_0^\beta d\tau\, H[\bar{\psi}_\alpha(\tau),\psi_\alpha(\tau)]\ , \tag{3.24b}$$

where

$$\psi_\alpha(0)=-\psi_\alpha(\beta)\ . \tag{3.24c}$$

This concludes the derivation of Eq. (3.1) and specifies the action.[4]

B. Sigma-model approach to disordered electronic systems

1. The model in matrix formulation

Wegner (1979) pioneered the formulation of the disordered fermion problem in terms of classical matrix fields. Wegner's original formulation started with bosonic (i.e., commuting) fields which describe diffusion of noninteracting electrons. Efetov *et al.* (1980) have given an equivalent formulation in terms of fermionic (i.e., Grassmannian) fields, and Finkel'stein (1983a) generalized the model to include interactions. While for noninteracting electrons the choice between a bosonic and a fermionic formulation is, to some extent, a matter of taste, the presence of interactions, and the resulting mixing of energies, requires the use of fermionic fields. In this section we present a generalized and expanded version of the model derivation given by Efetov *et al.* and by Finkel'stein.

a. The model in terms of Grassmann fields

We consider a system of fermions with an action

$$S=-\int dx\sum_\sigma\bar{\psi}_\sigma(x)\frac{\partial}{\partial\tau}\psi_\sigma(x)+S_0+S_{\rm dis}+S_{\rm int}\ . \tag{3.25a}$$

Here we use four-vector notation, $x\equiv(\mathbf{x},\tau)$, and $\int dx\equiv\int d\mathbf{x}\int_0^\beta d\tau$. S_0 describes free fermions with chemical potential μ,

$$S_0=\int dx\sum_\sigma\bar{\psi}_\sigma(x)\left[\frac{1}{2m}\Delta+\mu\right]\psi_\sigma(x)\ , \tag{3.25b}$$

with Δ the Laplace operator. $S_{\rm dis}$ describes non-spin-flip scattering by a random potential $u(\mathbf{x})$,

$$S_{\rm dis}=-\int dx\sum_\sigma\bar{\psi}_\sigma(x)u(\mathbf{x})\psi_\sigma(x)\ , \tag{3.25c}$$

and $S_{\rm int}$ describes a general two-particle interaction,

[4]This result (e.g., Negele and Orland, 1988) has been derived by ordering the imaginary-time slices in the usual way, from right to left. Some authors (e.g., Popov, 1983) give the result with a different sign of $\partial/\partial\tau$ in Eq. (3.24b). This one obtains by ordering the time slices from left to right. Since K is independent of imaginary time, time ordering is inessential, and the two results are equivalent.

$$S_{\rm int}=\frac{-1}{2}\int dx_1dx_2dx_3dx_4\sum_{\sigma_1,\sigma_2}v(x_1-x_2,x_3-x_2,x_4-x_1)\bar{\psi}_{\sigma_1}(x_1)\bar{\psi}_{\sigma_2}(x_2)\psi_{\sigma_2}(x_3)\psi_{\sigma_1}(x_4)\ . \tag{3.25d}$$

In writing Eq. (3.25d) we have used translational invariance in space and time. We assume the interaction potential v to be real and spin independent. Hermiticity requires

$$v(x,y,z)=v(x-y+z,-y,-z)\ . \tag{3.26}$$

We assume the random potential to be governed by a Gaussian distribution,

$$P[u]=\exp\left[-\pi N_F\tau\int d\mathbf{x}(u(\mathbf{x}))^2\right]\Big/\int D[u]\exp\left[-\pi N_F\tau\int d\mathbf{x}(u(\mathbf{x}))^2\right]\ . \tag{3.27}$$

The u-u correlation function then is

$$\{u(\mathbf{x})u(\mathbf{y})\}_{\rm dis}\equiv\int D[u]u(\mathbf{x})u(\mathbf{y})P[u]=\frac{1}{2\pi N_F\tau}\delta(\mathbf{x}-\mathbf{y})\ , \tag{3.28}$$

where τ is the elastic-scattering mean free time and N_F is the (bare) one-particle density of states.

Quenched (as opposed to annealed) disorder is characterized by static impurities and the necessity to average the free energy rather than the partition function (see, for example, Grinstein, 1985). This is accomplished by means of the replica trick (Edwards and Anderson, 1975), which is based on the identity

$$\ln Z=\lim_{N\to 0}\frac{1}{N}(Z^N-1)\ . \tag{3.29}$$

One introduces N identical "replicas" of the system (with N an integer) labeled by the index α. Then

$$Z^N=\int\prod_{\alpha=1}^{N}D[\bar{\psi}^\alpha,\psi^\alpha]\exp\left[\sum_{\alpha=1}^{N}S^\alpha[\bar{\psi}^\alpha,\psi^\alpha]\right]\ . \tag{3.30}$$

With a Gaussian distribution for the randomness, the calculation of the ensemble average $\{Z^N\}_{\rm dis}$ amounts simply to performing a Gaussian integral. The replica trick consists of performing this average for integer N, continuing the result analytically to real N, and taking the limit $N\to 0$.[5] Via Eq. (3.29), this procedure yields $\{\ln Z\}_{\rm dis}$. It reduces the calculation of a quenched average to that of an annealed average for the replicated system.

In order to calculate correlation functions, we add sources to the action,

$$S^\alpha\to S^\alpha+\sum_{i=1}^{\infty}\int\prod_{j=1}^{2i}dx_j\sum_{\{\sigma_i\}}J^{(2i)}_{\sigma_1,\ldots,\sigma_{2i}}(x_1,\ldots,x_{2i})\bar{\psi}^\alpha_{\sigma_1}(x_1)\cdots\bar{\psi}^\alpha_{\sigma_i}(x_i)\psi^\alpha_{\sigma_{i+1}}(x_{i+1})\cdots\psi^\alpha_{\sigma_{2i}}(x_{2i})\ , \tag{3.31}$$

and consider

$$\tilde{S}=\ln\int D[u]P[u]\exp\left[\sum_{\alpha=1}^{N}S^\alpha\right]\ . \tag{3.32}$$

We define correlation functions for the replicated theory as averages formed with the action $\tilde{S}$. These can be generated by differentiating with respect to the source fields $J^{(2i)}$. For instance, the two-point correlation function reads

$$\langle\bar{\psi}^\alpha_{\sigma_1}(x_1)\psi^\alpha_{\sigma_2}(x_2)\rangle\equiv\frac{1}{N\tilde{Z}}\int\prod_{\beta=1}^{N}D[\bar{\psi}^\beta,\psi^\beta]\left[\sum_{\alpha=1}^{N}\bar{\psi}^\alpha_{\sigma_1}(x_1)\psi^\alpha_{\sigma_2}(x_2)\right]\exp[\tilde{S}]$$

$$=\frac{1}{N}\frac{\delta}{\delta J^{(2)}_{\sigma_1\sigma_2}(x_1,x_2)}\ln\tilde{Z}\Big|_{J^{(2i)}=0}\ , \tag{3.33a}$$

[5]The mathematical legitimacy of this procedure has not been established for the model under consideration. This point is not entirely academic, as one can construct models for which the replica trick demonstrably fails (e.g., Verbaarschot and Zirnbauer, 1985). For the description of phase transitions an important question is whether or not the replica limit commutes with the bulk or thermodynamic limit. For noninteracting fermions, a way to avoid the replica trick is the supersymmetry method (Efetov, 1982). However, it comes with mathematical problems of its own and has not been applied to the interacting problem.

with a partition function

$$\tilde{Z} \equiv \int \prod_{\alpha=1}^{N} D[\bar{\psi}^{\alpha},\psi^{\alpha}]\exp[\tilde{S}] \ . \tag{3.33b}$$

We then obtain

$$\{\langle \bar{\psi}_{\sigma_1}(x_1)\psi_{\sigma_2}(x_2)\rangle\}_{\rm dis} = \lim_{N\to 0} \langle \bar{\psi}^{\alpha}_{\sigma_1}(x_1)\psi^{\alpha}_{\sigma_2}(x_2)\rangle \ , \tag{3.34}$$

where the correlation function on the left-hand side constitutes an average with the action S. The same result obviously applies to all higher correlation functions. We now have the following prescription. Replicate the system, and perform an annealed average. Calculate the correlation functions, and let $N\to 0$ ("replica limit"). This yields the averaged correlation functions of the original system with quenched disorder.

Let us define Fourier transforms of the Grassmann fields and the interaction potential,

$$\psi^{\alpha}_{\sigma}(x) = T^{1/2}\sum_{k} e^{i(\mathbf{k}\mathbf{x}-\omega_n\tau)}\psi^{\alpha}_{\sigma}(k) \ , \tag{3.35a}$$

$$\bar{\psi}^{\alpha}_{\sigma}(x) = T^{1/2}\sum_{k} e^{-i(\mathbf{k}\mathbf{x}-\omega_n\tau)}\bar{\psi}^{\alpha}_{\sigma}(k) \ , \tag{3.35b}$$

$$v(x_1,x_2,x_3) = T^3 \sum_{k_1,k_2,k_3} e^{-i\sum_{j=1}^{3} s_j(\mathbf{k}_j\mathbf{x}_j-\omega_{n_j}\tau_j)} \times v(k_1,k_2,k_3) \ , \tag{3.35c}$$

where $s_1=s_2=1$, $s_3=-1$, and for the spatial transform we have set the normalization volume $V=1$. $\omega_n = 2\pi T(n+1/2)$, $n=0,\pm 1,\ldots$, is a fermionic Matsubara frequency, and we use a four-vector notation $k\equiv(\mathbf{k},\omega_n)$. We now perform the impurity average according to Eq. (3.32). The Gaussian integration is easy,[6] and we obtain

$$\tilde{S} = \sum_{\alpha=1}^{n} (S^{\alpha}_0 + \tilde{S}^{\alpha}_{\rm dis} + S^{\alpha}_{\rm int}) \ , \tag{3.36a}$$

where

$$S^{\alpha}_0 = \sum_{\sigma}\sum_{k} \bar{\psi}^{\alpha}_{\sigma}(k)[i\omega_n - \mathbf{k}^2/2m + \mu]\psi^{\alpha}_{\sigma}(k) \ , \tag{3.36b}$$

$$\tilde{S}^{\alpha}_{\rm dis} = \frac{1}{4\pi N_F \tau}\sum_{\beta=1}^{N}\sum_{\{\mathbf{k}_i\}} \delta_{\mathbf{k}_1+\mathbf{k}_3,\mathbf{k}_2+\mathbf{k}_4} \times \sum_{\substack{n,m\\ \sigma,\sigma'}} \bar{\psi}^{\alpha}_{n\sigma}(\mathbf{k}_1)\psi^{\alpha}_{n\sigma}(\mathbf{k}_2)\bar{\psi}^{\beta}_{m\sigma'}(\mathbf{k}_3)\psi^{\beta}_{m\sigma'}(\mathbf{k}_4) \ , \tag{3.36c}$$

with $\psi_{n\sigma}(\mathbf{k}) = \psi_{\sigma}(\mathbf{k},\omega_n)$, etc., and

$$S^{\alpha}_{\rm int} = \frac{-T}{2}\sum_{\sigma,\sigma'}\sum_{\{k_i\}} \delta_{k_1+k_2,k_3+k_4} w(k_1,k_2,k_3) \times \bar{\psi}^{\alpha}_{\sigma}(k_1)\bar{\psi}^{\alpha}_{\sigma'}(k_2)\psi^{\alpha}_{\sigma'}(k_3)\psi^{\alpha}_{\sigma}(k_4) \ . \tag{3.36d}$$

The potential w is related to v by

$$w(k_1,k_2,k_3) = v(k_2-k_3,k_3,k_3-k_2-k_1) \tag{3.37a}$$

and obeys the following symmetry relations (which follow from Hermiticity):

$$\begin{aligned} w(k_1,k_2,k_3) &= w(k_1+k_2-k_3,k_3,k_2) \\ &= w(k_3,k_1+k_2-k_3,k_1) \\ &= w(k_2,k_1,k_1+k_2-k_3) \ . \end{aligned} \tag{3.37b}$$

Disorder-averaged correlation functions can be obtained from the action $\tilde{S}$, Eqs. (3.36), via Eqs. (3.33) and (3.34). In what follows we drop the tilde on $\tilde{S}$ and imply the replica limit.

b. Spinor notation and separation in phase space

In order to make the theory more tractable, we shall want to perform a Gaussian transformation to classical variables. To do so, it is convenient first to regroup the Grassmann variables. We follow Efetov *et al.* (1980) in defining bispinors,

$$\eta^{\alpha}_{n}(\mathbf{x}) = \frac{1}{\sqrt{2}}\begin{bmatrix} \bar{\psi}^{\alpha}_{n\uparrow}(\mathbf{x}) \\ \bar{\psi}^{\alpha}_{n\downarrow}(\mathbf{x}) \\ \psi^{\alpha}_{n\downarrow}(\mathbf{x}) \\ -\psi^{\alpha}_{n\uparrow}(\mathbf{x}) \end{bmatrix} . \tag{3.38}$$

The four degrees of freedom comprised by these objects are the charge or particle-hole degrees of freedom and the two spin degrees of freedom. For later reference, we define a basis in the space of complex 4×4 matrices as $\tau_i \otimes s_j$ $(i,j=0,1,2,3)$ with

$$\tau_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \tau_1 = \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix}, \quad \tau_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad \tau_3 = \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} \tag{3.39a}$$

and

$$s_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad s_1 = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}, \quad s_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad s_3 = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} . \tag{3.39b}$$

The τ_i are the quaternion basis and span the charge space, while the s_i serve as our basis in spin space. Notice that $\tau_j = -s_j = -i\sigma_j$ $(j=1,2,3)$, with σ_j the Pauli matrices. We introduce a spinor notation

[6] Gaussian integrals of the type encountered here can be performed by expanding the exponential and integrating term by term. In symbolic notation, $\int D[u]\exp[-(u^2+2u\bar{\psi}\psi)] = \int D[u]\exp[-u^2]\ [1-2u\bar{\psi}\psi+2u^2(\bar{\psi}\psi)^2+\cdots] = \text{const}\times\exp[(\bar{\psi}\psi)^2]$. The last identity can easily be proven by induction. The same result is obtained by completing the square and formally substituting $u\to u-\bar{\psi}\psi$.

$$\psi_n^\alpha(\mathbf{x})=\begin{bmatrix}\psi_{n\uparrow}^\alpha(\mathbf{x})\\ \psi_{n\downarrow}^\alpha(\mathbf{x})\end{bmatrix} . \tag{3.40}$$

Then we can write the bispinor as

$$\eta_n^\alpha(\mathbf{x})=\frac{1}{\sqrt{2}}\begin{bmatrix}\bar{\psi}_n^\alpha(\mathbf{x})\\ s_2\psi_n^\alpha(\mathbf{x})\end{bmatrix} \tag{3.41a}$$

and define an adjoint bispinor,

$$\bar{\eta}_n^\alpha(\mathbf{x})=(C\eta_n^\alpha(\mathbf{x}))^T=\frac{1}{\sqrt{2}}(-\psi_n^\alpha(\mathbf{x})^T,\bar{\psi}_n^\alpha(\mathbf{x})^T s_2^T) , \tag{3.41b}$$

with the charge-conjugation matrix

$$C=\begin{bmatrix}0 & s_2\\ s_2 & 0\end{bmatrix}=i\tau_1\otimes s_2 . \tag{3.41c}$$

We define Fourier transforms of the bispinors,

$$\begin{aligned}\eta^\alpha(k)&=\int d\mathbf{x}\, e^{i\mathbf{kx}}\eta_n^\alpha(\mathbf{x}) ,\\ \bar{\eta}^\alpha(k)&=\int d\mathbf{x}\, e^{-i\mathbf{kx}}\bar{\eta}_n^\alpha(\mathbf{x}) .\end{aligned} \tag{3.42}$$

Notice that $\eta(\mathbf{k},\omega_n)$ is composed of $\bar{\psi}(-\mathbf{k},\omega_n)$ and $\psi(\mathbf{k},\omega_n)$. With this notation, we can write the free-fermion part of the action as

$$S_0=\sum_\alpha\sum_k(\eta^\alpha(k),[i\omega_n-\mathbf{k}^2/2m+\mu]\eta^\alpha(k)) \tag{3.43}$$

and the disorder part as

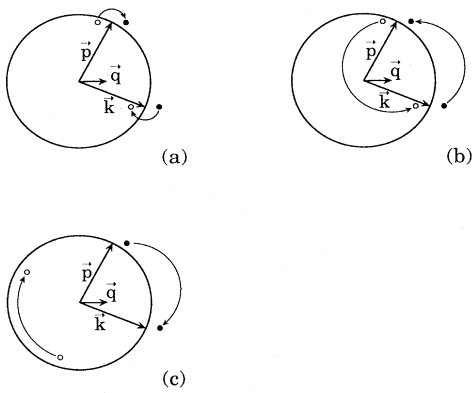

FIG. 12. Particle-hole scattering processes: (a) Small-angle scattering; (b) large-angle scattering; (c) $2k_F$ scattering of a particle-hole pair.

$$S_{\rm dis}=\frac{1}{4\pi N_F\tau}\sum_{\alpha,\beta}\sum_{n,m}\sum_{\{\mathbf{k}_i\}}\delta_{\mathbf{k}_1+\mathbf{k}_3,\mathbf{k}_2+\mathbf{k}_4}\times(\eta_n^\alpha(\mathbf{k}_1),\eta_n^\alpha(\mathbf{k}_2))(\eta_m^\beta(\mathbf{k}_3),\eta_m^\beta(\mathbf{k}_4)) . \tag{3.44}$$

Here $(\eta,\eta)\equiv\bar{\eta}\cdot\eta$ denotes a scalar product in bispinor space, which is given by the matrix product.

Before we turn to the interaction part of the action, let us consider the wave-number constraint in Eq. (3.44). $S_{\rm dis}$ can be rewritten in three different ways, which are identical,

$$\begin{aligned}4\pi N_F\tau S_{\rm dis}&=\sum_{\alpha,\beta}\sum_{n,m}\sum_{\mathbf{k},\mathbf{p},\mathbf{q}}(\eta_n^\alpha(\mathbf{k}),\eta_n^\alpha(\mathbf{k}+\mathbf{q}))(\eta_m^\beta(\mathbf{p}),\eta_m^\beta(\mathbf{p}-\mathbf{q}))\\ &=\sum_{\alpha,\beta}\sum_{n,m}\sum_{\mathbf{k},\mathbf{p},\mathbf{q}}(\eta_n^\alpha(\mathbf{k}),\eta_n^\alpha(\mathbf{p}))(\eta_m^\beta(\mathbf{p}+\mathbf{q}),\eta_m^\beta(\mathbf{k}+\mathbf{q}))\\ &=\sum_{\alpha,\beta}\sum_{n,m}\sum_{\mathbf{k},\mathbf{p},\mathbf{q}}(\eta_n^\alpha(\mathbf{k}),\eta_n^\alpha(\mathbf{p}))(\eta_m^\beta(-\mathbf{k}+\mathbf{q}),\eta_m^\beta(-\mathbf{p}+\mathbf{q})) .\end{aligned} \tag{3.45}$$

Suppose $\mathbf{q}$ is small, i.e., $|\mathbf{q}|\ll k_F$ with k_F the Fermi wave number. If we reexpress the η's in terms of the ψ's, we can then give a distinct physical interpretation to the three different terms that appear: small-angle scattering, large-angle scattering, and across-the-Fermi-surface or $2k_F$ scattering of a particle-hole pair by an impurity, respectively. This is illustrated in Fig. 12. Alternatively, Eq. (3.45) can be interpreted as taking into account different types of density fluctuations that are slowly decaying in the long-wavelength or $q\to 0$ limit. This is relevant because the aim of the field-theory approach is to identify and to construct an effective theory for the slow modes. Notice that the spinor representation mixes the second and third processes. Also notice that, since the sums in Eq. (3.45) are over *all* $\mathbf{q}$, each of the three different ways to write $S_{\rm dis}$ actually covers all possible processes. Our ultimate goal is to construct an effective theory for long-wavelength excitations. If one restricts the $\mathbf{q}$ sum in each of the three expressions in Eq. (3.45), each of them covers a different region in phase space. It has been universally assumed in the literature that it is permissible simply to add these three contributions. The second and the third are easily seen to be identical. We can thus write

$$S_{\rm dis}\simeq S_{\rm dis}^{(1)}+S_{\rm dis}^{(2)} , \tag{3.46a}$$

where

$$S_{\rm dis}^{(1)}=\frac{1}{4\pi N_F\tau}\sum_{\alpha,\beta}\sum_{n,m}\sum_{\mathbf{k},\mathbf{p}}\sum_{\mathbf{q}}{}'(\eta_n^\alpha(\mathbf{k}),\eta_n^\alpha(\mathbf{k}+\mathbf{q}))(\eta_m^\beta(\mathbf{p}),\eta_m^\beta(\mathbf{p}-\mathbf{q})) , \tag{3.46b}$$

$$S_{\rm dis}^{(2)}=\frac{1}{2\pi N_F\tau}\sum_{\alpha,\beta}\sum_{n,m}\sum_{\mathbf{k},\mathbf{p}}\sum_{\mathbf{q}}{}'(\eta_n^\alpha(\mathbf{k}),\eta_n^\alpha(\mathbf{p}))(\eta_m^\beta(\mathbf{p}+\mathbf{q}),\eta_m^\beta(\mathbf{k}+\mathbf{q})) . \tag{3.46c}$$

Here the prime on the q-summation symbol indicates the restriction to $|\mathbf{q}| \ll k_F$.

We now turn our attention to $S_{\rm int}$, Eq. (3.36d). Here it is advantageous, and physically more transparent, to perform the breakup into the three phase-space regions while still in the ψ representation. We thus write

$$S_{\rm int} \simeq S_{\rm int}^{(1)} + S_{\rm int}^{(2)} + S_{\rm int}^{(3)} \ , \tag{3.47a}$$

where

$$S_{\rm int}^{(1)} = \frac{-T}{2} \sum_{\alpha} \sum_{\sigma,\sigma'} \sum_{k,p} \sum_{q}{}' w(k,p+q,p) \bar{\psi}_{\sigma}^{\alpha}(k) \bar{\psi}_{\sigma'}^{\alpha}(p+q) \psi_{\sigma'}^{\alpha}(p) \psi_{\sigma}^{\alpha}(k+q) \ , \tag{3.47b}$$

$$S_{\rm int}^{(2)} = \frac{-T}{2} \sum_{\alpha} \sum_{\sigma,\sigma'} \sum_{k,p} \sum_{q}{}' w(k,p+q,k+q) \bar{\psi}_{\sigma}^{\alpha}(k) \bar{\psi}_{\sigma'}^{\alpha}(p+q) \psi_{\sigma'}^{\alpha}(k+q) \psi_{\sigma}^{\alpha}(p) \ , \tag{3.47c}$$

$$S_{\rm int}^{(3)} = \frac{-T}{2} \sum_{\alpha} \sum_{\sigma,\sigma'} \sum_{k,p} \sum_{q}{}' w(k,-k+q,p+q) \bar{\psi}_{\sigma}^{\alpha}(k) \bar{\psi}_{\sigma'}^{\alpha}(-k+q) \psi_{\sigma'}^{\alpha}(p+q) \psi_{\sigma}^{\alpha}(-p) \ . \tag{3.47d}$$

We next split these terms into spin-singlet and spin-triplet contributions, respectively. With the help of the spinor notation, Eq. (3.40), we can write

$$S_{\rm int}^{(1)} + S_{\rm int}^{(2)} = S_{\rm int}^{(s)} + S_{\rm int}^{(t)} \ , \tag{3.48a}$$

where

$$S_{\rm int}^{(s)} = \frac{-T}{2} \sum_{\alpha} \sum_{k,p} \sum_{q}{}' \Gamma_{k,p}^{(s)}(q) (\psi^{\alpha}(k), s_0 \psi^{\alpha}(k+q)) (\psi^{\alpha}(p+q), s_0 \psi^{\alpha}(p)) \ , \tag{3.48b}$$

$$S_{\rm int}^{(t)} = \frac{-T}{2} \sum_{\alpha} \sum_{k,p} \sum_{q}{}' \Gamma_{k,p}^{(t)}(q) \sum_{i=1}^{3} (\psi^{\alpha}(k), s_i \psi^{\alpha}(k+q)) (\psi^{\alpha}(p+q), s_i \psi^{\alpha}(p)) \ . \tag{3.48c}$$

Here we have defined singlet (s) and triplet (t) interaction amplitudes

$$\Gamma_{k,p}^{(t)}(q) = \tfrac{1}{2} w(k,p+q,k+q) \ , \tag{3.49a}$$

$$\Gamma_{k,p}^{(s)}(q) = w(k,p+q,p) - \Gamma_{k,p}^{(t)}(q) \ . \tag{3.49b}$$

They obey the symmetry relations

$$\Gamma_{k,p}^{(s,t)}(q) = \Gamma_{p,k}^{(s,t)}(q) = \Gamma_{k+q,p+q}^{(s,t)}(-q) \ . \tag{3.50}$$

$S_{\rm int}^{(s)}$ and $S_{\rm int}^{(t)}$ describe small frequency-momentum transfer between a particle and a hole. They are therefore referred to as the particle-hole interaction channel. $S_{\rm int}^{(3)}$, on the other hand, describes small energy-momentum transfer between two particles or two holes. It is therefore referred to as the particle-particle or Cooper interaction channel. Here the splitting into spin-singlet and spin-triplet contributions is straightforward, and we write

$$S_{\rm int}^{(3)} = S_{\rm int}^{c(s)} + S_{\rm int}^{c(t)} \ , \tag{3.51a}$$

where

$$S_{\rm int}^{c(s)} = -T \sum_{\alpha} \sum_{\sigma \neq \sigma'} \sum_{k,p} \sum_{q}{}' \Gamma_{k,p}^{c(s)}(q) \bar{\psi}_{\sigma}^{\alpha}(k) \bar{\psi}_{\sigma'}^{\alpha}(-k+q) \psi_{\sigma'}^{\alpha}(p+q) \psi_{\sigma}^{\alpha}(-p) \ , \tag{3.51b}$$

$$S_{\rm int}^{c(t)} = -T \sum_{\alpha} \sum_{\sigma,\sigma'} \sum_{k,p} \sum_{q}{}' \Gamma_{k,p}^{c(t)}(q) \bar{\psi}_{\sigma}^{\alpha}(k) \bar{\psi}_{\sigma'}^{\alpha}(-k+q) \psi_{\sigma'}^{\alpha}(p+q) \psi_{\sigma}^{\alpha}(-p) \ . \tag{3.51c}$$

The superscript c refers to the Cooper channel, and we have defined the amplitudes

$$\Gamma_{k,p}^{c(s,t)}(q) = \tfrac{1}{4} [w(k,-k+q,p+q) \pm w(k,-k+q,-p)] \ , \tag{3.52}$$

which obey the symmetry relations

$$\Gamma_{k,p}^{c(s,t)}(q) = \pm \Gamma_{k,-p-q}^{c(s,t)}(q) = \Gamma_{-p,-k}^{c(s,t)}(q) = \Gamma_{p+q,k-q}^{c(s,t)}(q) \ . \tag{3.53}$$

In Eqs. (3.52) and (3.53) the plus and minus signs apply to the singlet and triplet amplitudes, respectively. We note that for a pointlike, instantaneous interaction, $\Gamma^{c(t)}$ vanishes. This is a consequence of the Pauli principle.

Finally we express $S_{\rm int}$ in terms of bispinors. We shall be concerned only with even-parity potentials. From now on we shall restrict ourselves to this case, i.e., we assume that the interaction amplitudes are invariant under

$(\mathbf{k},\mathbf{p},\mathbf{q})\rightarrow(-\mathbf{k},-\mathbf{p},-\mathbf{q})$. Using this assumption as well as the symmetry properties (3.50) and (3.53) one easily convinces oneself that the four contributions to $S_{\rm int}$ can be written as

$$S_{\rm int}^{(s)}=-\frac{T}{2}\sum_{\alpha}\sum_{k,p}\sum_{q}{}'\sum_{r=0,3}(-)^r(\eta^\alpha(k),(\tau_r\otimes s_0)\eta^\alpha(k+q))\Gamma_{k,p}^{(s)}(q)(\eta^\alpha(p+q),(\tau_r\otimes s_0)\eta^\alpha(p))\ , \tag{3.54a}$$

$$S_{\rm int}^{(t)}=-\frac{T}{2}\sum_{\alpha}\sum_{k,p}\sum_{q}{}'\sum_{r=0,3}(-)^r\sum_{i=1}^{3}(\eta^\alpha(k),(\tau_r\otimes s_i)\eta^\alpha(k+q))\Gamma_{k,p}^{(t)}(q)(\eta^\alpha(p+q),(\tau_r\otimes s_i)\eta^\alpha(p))\ , \tag{3.54b}$$

$$S_{\rm int}^{c(s)}=-\frac{T}{2}\sum_{\alpha}\sum_{\mathbf{k},\mathbf{p}}\sum_{\mathbf{q}}{}'\sum_{n_1,m;n_2}\sum_{r=1,2}(\eta^\alpha_{n_1}(-\mathbf{k}),(\tau_r\otimes s_0)\eta^\alpha_{-n_1+m}(-\mathbf{k}+\mathbf{q}))\Gamma^{c(s)}_{\substack{\mathbf{k},\mathbf{p}\\ n_1,n_2}}(\mathbf{q},m)$$
$$\times(\eta^\alpha_{-n_2}(-\mathbf{p}),(\tau_r\otimes s_0)\eta^\alpha_{n_2+m}(-\mathbf{p}-\mathbf{q}))\ , \tag{3.54c}$$

$$S_{\rm int}^{c(t)}=-\frac{T}{2}\sum_{\alpha}\sum_{\mathbf{k},\mathbf{p}}\sum_{\mathbf{q}}{}'\sum_{n_1,m;n_2}\sum_{r=1,2}\sum_{i=1}^{3}(\eta^\alpha_{n_1}(-\mathbf{k}),(\tau_r\otimes s_i)\eta^\alpha_{-n_1+m}(-\mathbf{k}+\mathbf{q}))$$
$$\times\Gamma^{c(t)}_{\substack{\mathbf{k},\mathbf{p}\\ n_1,n_2}}(\mathbf{q},m)(\eta^\alpha_{-n_2}(-\mathbf{p}),(\tau_r\otimes s_i)\eta^\alpha_{n_2+m}(-\mathbf{p}-\mathbf{q}))\ . \tag{3.54d}$$

c. Decoupling of the four-fermion terms

The next step is to perform a Gaussian transformation (Stratonovich, 1957; Hubbard, 1959) on the action. The objective is to integrate out the Grassmann fields and to formulate the theory in terms of c-number fields instead. Among other things, the resulting field theory can then be examined or approximately solved using saddle-point techniques.

We start with $S_{\rm dis}^{(1)}$, Eq. (3.46b). If we define a Grassmannian field

$$A(\mathbf{q})=\frac{1}{\sqrt{4\pi N_F\tau}}\sum_{\alpha}\sum_{n}\sum_{\mathbf{k}}(\eta^\alpha_n(\mathbf{k}),\eta^\alpha_n(\mathbf{k}+\mathbf{q}))\ , \tag{3.55}$$

then we can write

$$S_{\rm dis}^{(1)}=\int d\mathbf{x}(A(\mathbf{x}))^2\ . \tag{3.56}$$

Here, and in all following real-space integrals, the restriction to long-wavelength fluctuations will be implicitly understood. Notice that $A(\mathbf{x})$ is self-adjoint, $\overline{A}(\mathbf{x})=A(\mathbf{x})$. Now we introduce a scalar c-number field $\nu(\mathbf{x})$. Then we can write the contribution of $S_{\rm dis}^{(1)}$ to the partition function as a Gaussian integral over $\nu(\mathbf{x})$,

$$\exp[S_{\rm dis}^{(1)}]=\int D[\nu]e^{-\int d\mathbf{x}[(\nu(\mathbf{x}))^2+2\nu(\mathbf{x})A(\mathbf{x})]}\ , \tag{3.57}$$

where we have again dropped a multiplicative constant. Since $A(\mathbf{x})$ is self-adjoint, $\nu(\mathbf{x})$ can be chosen to be real.

Now consider $S_{\rm dis}^{(2)}$. Here the terms that belong together in $\mathbf{k}$ space are members of different scalar products, so we shall have to work with matrix fields. We introduce covariant and contravariant indices such that ${}^i\eta$ denotes the elements of η, and ${}_i\eta$ denotes the elements of $\overline{\eta}$. If we define a Grassmannian matrix field

$${}_i^jB^{\alpha\beta}_{nm}(\mathbf{q})=\frac{1}{\sqrt{2\pi N_F\tau}}\sum_{\mathbf{k}}{}_i\eta^\alpha_n(\mathbf{k})^j\eta^\beta_m(\mathbf{k}+\mathbf{q})\ , \tag{3.58}$$

then we can write

$$S_{\rm dis}^{(2)}=-\sum_{\alpha,\beta}\sum_{n,m}\sum_{\mathbf{q}}{}'\sum_{i,j}{}_i{}^jB^{\alpha\beta}_{nm}(\mathbf{q})_j{}^iB^{\beta\alpha}_{mn}(-\mathbf{q})$$
$$=-\int d\mathbf{x}\,\mathrm{tr}(B(\mathbf{x}))^2\ . \tag{3.59}$$

Here tr denotes the trace over all discrete indices. B has the property $\overline{B}(\mathbf{x})=B^T(\mathbf{x})$. Now we introduce a c-number matrix field $Q(\mathbf{x})$. Then we can write

$$\exp[S_{\rm dis}^{(2)}]=\int D[Q]\exp\left\{-\int d\mathbf{x}\left[\mathrm{tr}(Q(\mathbf{x}))^2+\frac{2i}{\sqrt{2\pi N_F\tau}}(\eta(\mathbf{x})|Q(\mathbf{x})\eta(\mathbf{x}))\right]\right\}\ . \tag{3.60}$$

Here we have introduced a new scalar product in bispinor space,

$$(\eta|\eta)=\sum_{\alpha}\sum_{n}\sum_{i}{}_i\eta^{\alpha i}_n\eta^\alpha_n\ . \tag{3.61}$$

Q is an infinite matrix whose matrix elements $Q^{\alpha\beta}_{nm}$ are spin quaternions and can be represented by complex 4×4 matrices. For instance, we can expand $Q^{\alpha\beta}_{nm}$ in the basis given by Eq. (3.39),

$$Q^{\alpha\beta}_{nm}(\mathbf{x})=\sum_{r=0}^{3}\sum_{i=0}^{3}{}^i_rQ^{\alpha\beta}_{nm}(\mathbf{x})(\tau_r\otimes s_i)\ . \tag{3.62}$$

Since $\overline{B}=B^T$, we can choose Q to be Hermitian, $Q^+=Q$.

Furthermore, as an operator in bispinor space, Q must respect the scalar product, Eq. (3.61). Since $\bar{\eta}$ and η are related by charge conjugation, Eq. (3.41b), this imposes the condition

$$Q^{+}=C^{T}Q^{T}C=Q \ . \tag{3.63}$$

From the first equality in Eq. (3.63) it follows that the elements of Q have the structure

$$Q^{\alpha\beta}_{nm}=\begin{bmatrix} d_{\uparrow\uparrow} & d_{\uparrow\downarrow} & -c_{\uparrow\downarrow} & c_{\uparrow\uparrow} \\ d_{\downarrow\uparrow} & d_{\downarrow\downarrow} & -c_{\downarrow\downarrow} & c_{\downarrow\uparrow} \\ c_{\downarrow\uparrow}^{*} & c_{\downarrow\downarrow}^{*} & d_{\downarrow\downarrow}^{*} & -d_{\downarrow\uparrow}^{*} \\ -c_{\uparrow\uparrow}^{*} & -c_{\uparrow\downarrow}^{*} & -d_{\uparrow\downarrow}^{*} & d_{\uparrow\uparrow}^{*} \end{bmatrix}^{\alpha\beta}_{nm} \tag{3.64}$$

The equivalent statement in the expansion (3.62) is that the elements of Q describing the particle-hole degrees of freedom are real, while those describing the particle-particle degrees of freedom are purely imaginary,

$${}^{i}_{r}Q^{\alpha\beta\,*}_{nm}={}^{i}_{r}Q^{\alpha\beta}_{nm}, \quad r=0,3 \ , \tag{3.65a}$$

$${}^{i}_{r}Q^{\alpha\beta\,*}_{nm}=-{}^{i}_{r}Q^{\alpha\beta}_{nm}, \quad r=1,2 \ . \tag{3.65b}$$

In addition, from the Hermiticity requirement we obtain the symmetry properties

$${}^{0}_{r}Q^{\alpha\beta}_{nm}=(-)^{r}\,{}^{0}_{r}Q^{\beta\alpha}_{mn}, \quad r=0,3 \ , \tag{3.66a}$$

$${}^{i}_{r}Q^{\alpha\beta}_{nm}=(-)^{r+1}\,{}^{i}_{r}Q^{\beta\alpha}_{mn}, \quad r=0,3;\ i=1,2,3 \ , \tag{3.66b}$$

$${}^{0}_{r}Q^{\alpha\beta}_{nm}={}^{0}_{r}Q^{\beta\alpha}_{mn}, \quad r=1,2 \ , \tag{3.66c}$$

$${}^{i}_{r}Q^{\alpha\beta}_{nm}=-{}^{i}_{r}Q^{\beta\alpha}_{mn}, \quad r=1,2;\ i=1,2,3 \ . \tag{3.66d}$$

Next we separate $S_{\rm int}$ by means of Gaussian transformations. In order to facilitate the calculations, we make two simplifying assumptions concerning the interaction amplitudes. First, we neglect the $\mathbf{k}$ and $\mathbf{p}$ dependence of the $\Gamma_{\mathbf{kp}}(\mathbf{q})$. For our purposes this does not mean any loss of generality. Later we shall construct an effective theory for slow (i.e., density) modes only. Therefore, we shall be interested only in the scalar component of the $\Gamma_{\mathbf{k},\mathbf{p}}$ in a multipole expansion. Second, we neglect the frequency dependence of the amplitudes $\Gamma^{(s,t)}$ and $\Gamma^{c(s)}$. This is consistent with the symmetry relations (3.50) and (3.53); all work done to date has been for instantaneous interactions. For $\Gamma^{c(t)}$, we cannot drop the frequency dependence lest the amplitude vanish, so in this channel we keep the full frequency dependence.

We define Grassmannian fields

$${}^{i}_{r}\Delta^{\alpha\beta}_{nm}(\mathbf{q})=\sum_{\mathbf{k}}(\eta^{\alpha}_{n}(\mathbf{k}),(\tau_{r}\otimes s_{i})\eta^{\beta}_{m}(\mathbf{k}+\mathbf{q})) \tag{3.67}$$

and four matrices $M^{(u)}$ $(u=1,2,3,4)$,

$${}^{i_1 i_2}_{r_1 r_2}M^{(u)\alpha_1\beta_1,\alpha_2\beta_2}_{n_1 m_1,n_2 m_2}(\mathbf{x},\mathbf{y})=\frac{-T}{2}\delta_{\alpha_1\beta_1}\delta_{\alpha_2\beta_2}\delta_{\alpha_1\alpha_2}\delta_{r_1 r_2}(-)^{r_1}[\delta_{r_1 0}+\delta_{r_1 3}]\delta_{i_1 i_2}\delta_{n_1+n_2,m_1+m_2}\Gamma^{(u)}(\mathbf{x}-\mathbf{y})\begin{cases}\delta_{i_1 0} & \text{for } u=1,\\ \sum_{j=1}^{3}\delta_{i_1 j} & \text{for } u=2,\end{cases} \tag{3.68a}$$

$${}^{i_1 i_2}_{r_1 r_2}M^{(u)\alpha_1\beta_1,\alpha_2\beta_2}_{n_1 m_1,n_2 m_2}(\mathbf{x},\mathbf{y})=\frac{-T}{2}\delta_{\alpha_1\beta_1}\delta_{\alpha_2\beta_2}\delta_{\alpha_1\alpha_2}\delta_{r_1 r_2}[\delta_{r_1 1}+\delta_{r_1 2}]\delta_{i_1 i_2}\delta_{n_1+m_1,n_2+m_2}\begin{cases}\Gamma^{(u)}(\mathbf{x}-\mathbf{y})\delta_{i_1 0} & \text{for } u=3,\\ \Gamma^{(u)}_{n_1,-n_2}(\mathbf{x}-\mathbf{y},m_1+n_1)\sum_{j=1}^{3}\delta_{i_1 j} & \text{for } u=4,\end{cases} \tag{3.68b}$$

where we have defined $\Gamma^{(1)}=\Gamma^{(s)}$, $\Gamma^{(2)}=\Gamma^{(t)}$, $\Gamma^{(3)}=\Gamma^{c(s)}$, $\Gamma^{(4)}=\Gamma^{c(t)}$. Then we can write

$$S_{\rm int}=\int d\mathbf{x}\,d\mathbf{y}\sum_{u=1}^{4}\sum_{1,2}\Delta_{1}(\mathbf{x})M^{(u)}_{12}(\mathbf{x}-\mathbf{y})\Delta_{2}(\mathbf{y}) \ , \tag{3.69}$$

where $1\equiv(\alpha_1\beta_1,r_1 i_1,n_1 m_1)$, etc. If we introduce a c-number field $X_1(\mathbf{x})$, then we can write

$$\exp[S_{\rm int}]=\int D[X]\exp\left[-\sum_{u=1}^{4}\int d\mathbf{x}\,d\mathbf{y}\sum_{1,2}X_{1}(\mathbf{x})M^{(u)}_{12}(\mathbf{x}-\mathbf{y})X_{2}(\mathbf{y})\right]\exp\left[-2\sum_{u=1}^{4}\int d\mathbf{x}(\eta(\mathbf{x})|0^{(u)}(\mathbf{x})\eta(\mathbf{x}))\right] \ . \tag{3.70}$$

Here we have used the fact that $M^{(u)}_{21}(\mathbf{y}-\mathbf{x})=M^{(u)}_{12}(\mathbf{x}-\mathbf{y})$, which follows from the definition of the $M^{(u)}$ and the symmetry relations for the Γ's. We also have defined four new operators in bispinor space

$$O^{(u)\alpha\beta}_{nm}(\mathbf{x})=\sum_{r}\sum_{i}{}^{i}_{r}N^{(u)\alpha\beta}_{nm}(\mathbf{x})\tau_{r}\otimes s_{i} \ , \tag{3.71a}$$

where

$$N^{(u)}_{1}(\mathbf{x})=\int d\mathbf{y}\sum_{2}M^{(u)}_{12}(\mathbf{x}-\mathbf{y})X_{2}(\mathbf{y}) \ . \tag{3.71b}$$

Let us now inspect Eqs. (3.43), (3.57), (3.60), (3.70), and (3.71). We see that we have rewritten the action in a form that is bilinear in η, at the expense of introducing new fields. The partition function can be written

$$Z=\int D[Q]\exp(S[Q])\ , \tag{3.72}$$

with

$$S[Q]=\ln\int D[\eta]D[\nu]D[X]e^{S_0[\bar{\eta},\eta]+S_{\rm dis}[Q,\nu,\bar{\eta},\eta]+S_{\rm int}[X,\bar{\eta},\eta]}\ . \tag{3.73}$$

The mentioned bilinearity of S in η allows one to integrate out the Grassmann fields by means of the operator identity

$$\int D[\eta]e^{\int d\mathbf{x}(\eta(\mathbf{x})|O\eta(\mathbf{x}))}=(\det O)^{1/2}=e^{\int d\mathbf{x}\,\mathrm{tr}(\ln O(\mathbf{x}))/2}\ , \tag{3.74}$$

which follows from Eq. (3.13a). The result can be simplified by shifting Q,

$$Q(\mathbf{x})\to Q(\mathbf{x})+i\nu(\mathbf{x})/\sqrt{2}+i\sqrt{2\pi N_F\tau}\sum_{u=1}^{4}O^{(u)}(\mathbf{x})\ . \tag{3.75}$$

$\nu(\mathbf{x})$ can then immediately be integrated out. Again we drop multiplicative constant contributions to Z. If we further notice that $\mathrm{tr}O^{(u)}\propto\delta_{u1}$, we can write the action in the form

$$\exp(S[Q])=\exp\left[-\int d\mathbf{x}\left[\mathrm{tr}[Q(\mathbf{x})]^2-\frac{1}{2}\mathrm{tr}\ln[i\omega_n+\Delta_{\mathbf{x}}/2m+\mu-2iQ(\mathbf{x})/\sqrt{2\pi N_F\tau}]\right]\right.$$
$$\left.-(1/2N_s)(\mathrm{tr}Q(\mathbf{x}))^2\right]\int D[X]\exp(-\mathcal{L}[X,Q])\ , \tag{3.76a}$$

where

$$\mathcal{L}[X,Q]=\sum_{u=1}^{4}\int d\mathbf{x}\,d\mathbf{y}\sum_{1,2}X_1(\mathbf{x})M_{12}^{(u)}(\mathbf{x}-\mathbf{y})X_2(\mathbf{y})$$
$$-2\pi N_F\tau\int d\mathbf{x}\,\mathrm{tr}\left[\sum_{u=1}^{4}O^{(u)}(\mathbf{x})\right]^2+2i\sqrt{2\pi N_F\tau}\int d\mathbf{x}\sum_{u=1}^{4}\mathrm{tr}(Q(\mathbf{x})O^{(u)}(\mathbf{x}))$$
$$+i\sqrt{2\pi N_F\tau}/N_s\int d\mathbf{x}\,\mathrm{tr}Q(\mathbf{x})\mathrm{tr}O^{(1)}(\mathbf{x})-(\pi N_F\tau/N_s)\int d\mathbf{x}(\mathrm{tr}O^{(1)}(\mathbf{x}))^2\ , \tag{3.76b}$$

and N_s contains a free sum over frequency indices,

$$N_s=1-2N\sum_{n}\ . \tag{3.76c}$$

We note that the resulting action is a quadratic form in X, so we can also integrate out X exactly. This procedure is more involved, and we devote the next subsection to it.

d. The model in terms of classical matrix fields

We simplify our notation by defining a scalar product in the space of the fields X_1 as

$$(X|X)=\int d\mathbf{x}\sum_{1}X_1(\mathbf{x})X_1(\mathbf{x})\ . \tag{3.77}$$

Then the first term in $\mathcal{L}$, Eq. (3.76b), is simply

$$\sum_{u=1}^{4}(X|M^{(u)}X)\ . \tag{3.78}$$

For the second term, we notice that $\mathrm{tr}O^{(u)}O^{(v)}\sim\delta_{uv}$ due to the spin and charge structure of the matrices $M^{(u)}$, Eqs. (3.68). It is therefore sufficient to consider

$$\int d\mathbf{x}\,\mathrm{tr}(O^{(u)}(\mathbf{x}))^2=(-)^u2T\Gamma^{(u)}\left[\sum_{n}\right](X|M^{(u)}X),\quad u=1,2\ , \tag{3.79a}$$

$$\int d\mathbf{x}\,\mathrm{tr}(O^{(3)}(\mathbf{x}))^2 = 2T\Gamma^{(3)}\left[\sum_n\right](X|M^{(3)}X)\ , \tag{3.79b}$$

$$\int d\mathbf{x}\,\mathrm{tr}(O^{(4)}(\mathbf{x}))^2 = -2T(X|P^{(4)}X)\ . \tag{3.79c}$$

In Eqs. (3.79a) and (3.79b) we have for simplicity assumed short-range interactions,

$$\Gamma^{(u)}(\mathbf{x}) = \Gamma^{(u)}\delta(\mathbf{x})\ . \tag{3.80a}$$

For our purposes this will not mean any loss of generality for $u=2,3$. On the other hand, the long-range nature of the Coulomb interaction implies that $\Gamma^{(1)}$ has a contribution that is not short ranged. This contribution will be considered separately in Sec. III.B.3.d below.

$P^{(4)}$ is defined as

$$P_{12}^{(4)}(\mathbf{x}) = \gamma_{12}^{(4)}(\mathbf{x})M_{12}^{(4)}(\mathbf{x})\ , \tag{3.80b}$$

with

$$\gamma_{12}^{(4)}(\mathbf{x}) = \gamma^{(4)}_{n_1m_1,n_2m_2}(\mathbf{x}) = \int d\mathbf{y}\sum_m \Gamma^{(4)}_{m_1,-m}(\mathbf{y},n_1+m_1)\Gamma^{(4)}_{m,-n_2}(\mathbf{y}+\mathbf{x},n_2+m_2)/\Gamma^{(4)}_{n_1,-n_2}(\mathbf{x},n_1+m_1)\ . \tag{3.80c}$$

Notice that, as a consequence of our assuming instantaneous interactions, free frequency sums appear in Eqs. (3.79a) and (3.79b). These sums, which formally are infinite (or linearly dependent on any frequency cutoff), will be of no consequence for observable qualities, as we shall see. For the next term in Eq. (3.76b), we find

$$\int d\mathbf{x}\,\mathrm{tr}(Q(\mathbf{x})O^{(u)}(\mathbf{x})) = \pm 4(Q|M^{(u)}X),\quad +\ \text{for } u=1,2;\ -\ \text{for } u=3,4\ . \tag{3.81}$$

Finally, we have to consider $\mathrm{tr}O^{(1)}(\mathbf{x})$. If we define a matrix

$$L_{12}(\mathbf{x}) = (4\pi/N_s)N_F\tau\left[T\Gamma^{(1)}\sum_n\right]^2\delta(\mathbf{x})\delta_{n_1m_1}\delta_{n_2m_2}\delta_{r_1r_2}\delta_{r_10}\delta_{i_1i_2}\delta_{i_10}\delta_{\alpha_1\beta_1}\delta_{\alpha_2\beta_2}\ , \tag{3.82}$$

then we can write

$$\int d\mathbf{x}(\mathrm{tr}O^{(1)}(\mathbf{x}))^2 = \frac{N_s}{\pi N_F\tau}(X|LX)\ , \tag{3.83a}$$

and

$$\int d\mathbf{x}\,\mathrm{tr}Q(\mathbf{x})\mathrm{tr}O^{(1)}(\mathbf{x}) = \frac{-2N_s/\pi N_F\tau}{T\Gamma^{(1)}\sum_n}(X|LQ)\ . \tag{3.83b}$$

Collecting our results, we now have

$$\mathcal{L}[X,Q] = \sum_{u=1}^{4}(X|\tilde{M}^{(u)}X) + 8i\sqrt{2\pi N_F\tau}\left[\sum_{u=1}^{2}(Q|M^{(u)}X) - \sum_{u=3}^{4}(Q|M^{(u)}X)\right] - (X|LX) - \frac{2i\sqrt{2\pi N_F\tau}}{\pi N_F\Gamma^{(1)}T\tau\sum_n}(X|LQ)\ , \tag{3.84}$$

where

$$\tilde{M}_{12}^{(u)}(\mathbf{x}) = \left[1 \pm 4\pi N_F\Gamma^{(u)}T\tau\sum_n\right]M_{12}^{(u)}(\mathbf{x}),\quad +\ \text{for } u=1,\ -\ \text{for } u=2,3\ , \tag{3.85a}$$

$$\tilde{M}_{12}^{(4)}(\mathbf{x}) = M_{12}^{(4)}(\mathbf{x}) + 4\pi T\tau N_F P_{12}^{(4)}(\mathbf{x})\ . \tag{3.85b}$$

Before we proceed to integrate out the fields X, let us pause and consider the structure of Eqs. (3.84) and (3.85). The operator L couples only to the particle-hole spin-singlet degree of freedom, and the matrices $\tilde{M}^{(1,2,3)}$ are simply multiples of the matrix M. The fields X, which describe the particle-hole triplet and the particle-particle singlet channels, respectively, can therefore be integrated out immediately. The particle-hole singlet channel, $u=1$, we shall deal with below. The result for the particle-particle spin-triplet channel, $u=4$, can be formally expressed in terms of the inverse of the operator $\tilde{M}^{(4)}$. However, this inverse is rather awkward to deal with in explicit calculations later. To deal with this problem, we note that the difference between $\tilde{M}^{(u)}$ and $M^{(u)}$ is of first order in the small parameter $T\tau$ for all u. Finkel'stein (1983a) has argued that these terms should therefore be irrelevant. We follow this reasoning and re-

place all $\tilde{M}$'s by M's.[7]

Integrating out X for the channels $u=2,3,4$ is now straightforward, and the contribution to the action is

$$S_{\rm int}^{(u=2,3,4)}[Q]=-32\pi N_F\tau\sum_{u=2}^{4}(Q|M^{(u)}|Q)\ . \tag{3.86}$$

For the particle-hole spin-singlet channel (i.e., $u=1$), we have

$$\begin{aligned}\mathcal{L}^{(u=1)}[X,Q]&=(X|M^{(1)}-L|X)-(Q|aM^{(1)}-bL|X)\\&\quad-(X|aM^{(1)}-bL|Q)\\&=(X-\tilde{Q}|M^{(1)}-L|X-\tilde{Q})\\&\quad-(\tilde{Q}|M^{(1)}-L|\tilde{Q})\ ,\end{aligned} \tag{3.87a}$$

with

$$\tilde{Q}=(M^{(1)}-L)^{-1}(aM^{(1)}-bL)Q \tag{3.87b}$$

and

$$a=-4i\sqrt{2\pi N_F\tau}\ , \tag{3.87c}$$

$$b=-i\sqrt{2\pi N_F\tau}/\pi N_F\Gamma^{(1)}T\tau\sum_n\ . \tag{3.87d}$$

The corresponding contribution to the action is

$$\begin{aligned}S_{\rm int}^{(u=1)}[Q]&=(Q|(aM^{(1)}-bL)(M^{(1)}-L)^{-1}(aM^{(1)}-bL)|Q)\\&=a^2(Q|M^{(1)}|Q)-b^2(Q|L|Q)+(a-b)^2(Q|[1-L(M^{(1)})^{-1}]^{-1}L|Q)\ ,\end{aligned} \tag{3.88}$$

So far we have assumed that the inverses of $M^{(1)}$ and L exist. This is not the case, since the operators

$$M_{n_1n_2,n_3n_4}=m\delta_{n_1+n_3,n_2+n_4}\ , \tag{3.89a}$$

$$L_{n_1n_2,n_3n_4}=l\delta_{n_1n_2}\delta_{n_3n_4} \tag{3.89b}$$

do not have an inverse, and neither does their sum. We can circumvent this problem by adding a small diagonal term to M,

$$M^{\varepsilon}_{n_1n_2,n_3n_4}=\varepsilon\delta_{n_1n_3}\delta_{n_2n_4}+M_{n_1n_2,n_3n_4}\ , \tag{3.89c}$$

and letting $\varepsilon\to 0$ in the end. M^{ε} is invertible, and we find

$$\begin{aligned}S_{\rm int}^{(u=1)}[Q]&=a^2(Q|M^{(1)}|Q)\\&\quad-\left[b^2+(a-b)^2\frac{m}{Nl-m}\right](Q|L|Q)\ .\end{aligned} \tag{3.90}$$

Here a and b are given by Eqs. (3.87c) and (3.87d); from Eqs. (3.68a) and (3.82) we have

$$m=-\frac{T}{2}\Gamma^{(1)}\ , \tag{3.91a}$$

$$l=4\left[T\Gamma^{(1)}\sum_n\right]^2\pi N_F\tau/N_s\ . \tag{3.91b}$$

Let us collect our results. With the help of Eqs. (3.76a), (3.86), and (3.90) we can write the action

$$\begin{aligned}S[Q]&=\frac{-\pi N_F}{8\tau}\int d\mathbf{x}\,{\rm tr}(Q(\mathbf{x}))^2\\&\quad+\int d\mathbf{x}\,{\rm tr}\ln[i\omega-H_0+iQ(\mathbf{x})/2\tau]/2+S_{\rm int}[Q]\ ,\end{aligned} \tag{3.92a}$$

where

$$H_0=-\frac{1}{2m}\Delta_{\mathbf{x}}-\mu \tag{3.92b}$$

and

$$S_{\rm int}[Q]=-\frac{\pi T}{4}\sum_{u=1}^{4}\int d\mathbf{x}[Q(\mathbf{x})\gamma^{(u)}Q(\mathbf{x})]+\delta S_{\rm int}[Q]\ , \tag{3.92c}$$

with

$$[Q\gamma^{(1)}Q]=K^{(1)}\sum_{\substack{n_1n_2\\n_3n_4}}\delta_{n_1+n_3,n_2+n_4}\sum_{\alpha}\sum_{r=0,3}(-)^r{\rm tr}(\tau_r\otimes s_0Q^{\alpha\alpha}_{n_1n_2}){\rm tr}(\tau_r\otimes s_0Q^{\alpha\alpha}_{n_3n_4})\ , \tag{3.92d}$$

$$[Q\gamma^{(2)}Q]=-K^{(2)}\sum_{\substack{n_1n_2\\n_3n_4}}\delta_{n_1+n_3,n_2+n_4}\sum_{\alpha}\sum_{r=0,3}(-)^r\sum_{i=1}^{3}{\rm tr}(\tau_r\otimes s_iQ^{\alpha\alpha}_{n_1n_2}){\rm tr}(\tau_r\otimes s_iQ^{\alpha\alpha}_{n_3n_4})\ , \tag{3.92e}$$

[7]The validity of this procedure is not obvious, mainly because of the infinite quantity $\sum_n$ multiplying the small parameter $T\tau$. For frequency-independent interaction amplitudes, one can actually keep the terms of $O(T\tau)$ and demonstrate that they are of no consequence. This is shown in Sec. III.B.2 below. However, by dropping the frequency dependence of $\Gamma^{(4)}$ one loses the particle-particle spin-triplet channel altogether. See the discussion after Eqs. (3.66).

$$[Q\gamma^{(3)}Q]=-K^{(3)}\sum_{\substack{n_1n_2\\n_3n_4}}\delta_{n_1+n_2,n_3+n_4}\sum_{\alpha}\sum_{r=1,2}\mathrm{tr}_s\{\mathrm{tr}_\tau(\tau_r\otimes s_0Q^{\alpha\alpha}_{n_1n_2})\mathrm{tr}_\tau(\tau_r\otimes s_0Q^{\alpha\alpha}_{n_3n_4})\}\ , \tag{3.92f}$$

$$[Q\gamma^{(4)}Q]=\sum_{\substack{n_1n_2\\n_3n_4}}K^{(4)}_{n_1n_2,n_3n_4}\delta_{n_1+n_2,n_3+n_4}\sum_{\alpha}\sum_{r=1,2}\sum_{i=1}^{3}\mathrm{tr}_s\{\mathrm{tr}_\tau(\tau_r\otimes s_iQ^{\alpha\alpha}_{n_1n_2})\mathrm{tr}_\tau(\tau_r\otimes s_iQ^{\alpha\alpha}_{n_3n_4})\}\ . \tag{3.92g}$$

Here $\mathrm{tr}=\mathrm{tr}_\tau\mathrm{tr}_s$ with tr_τ acting on the τ's and tr_s acting on the s's, and we have defined new interaction amplitudes

$$K^{(u)}=\begin{cases}(-1)^u\pi N_F^2\Gamma^{(u)}/2, & u=1,2,\\ \pi N_F^2\Gamma^{(u)}, & u=3,4.\end{cases} \tag{3.92h}$$

In later sections we shall go back to the notation $K^{(s)}$, $K^{(t)}$, $K^{c(s)}$, and $K^{c(t)}$ instead of $K^{(1)}$, $K^{(2)}$, $K^{(3)}$, and $K^{(4)}$, respectively. Note that $K^{(s)}<0$ for a repulsive interaction. Finally, δS_{int} in Eq. (3.92c) is given by

$$\delta S_{\mathrm{int}}[Q]=-\left[b^2+(a-b)^2\frac{m}{Nl-m}\right]\frac{\pi^2}{32N_s}$$
$$\times N_F^2\left[T\Gamma^{(1)}\sum_n\right]^2\int d\mathbf{x}(\mathrm{tr}Q(\mathbf{x}))^2\ , \tag{3.92i}$$

with a, b, l, and m given by Eqs. (3.87c), (3.87d), (3.91a), (3.91b), and N the number of replicas. Notice that in writing Eqs. (3.92) we have scaled the Q fields with a factor $-\sqrt{\pi N_F/8\tau}$.

2. Saddle-point solutions

In this section we consider solutions of the field theory given by Eqs. (3.92) which one obtains in the saddle-point approximation. The saddle-point equation is

$$\delta S[Q]/\delta^i_rQ^{\alpha\beta}_{nm}(\mathbf{x})=0\ . \tag{3.93}$$

The saddle-point solution of the noninteracting theory, Eq. (3.92a) with $S_{\mathrm{int}}=0$, has served as a starting point for the derivation of the nonlinear sigma model, both for the noninteracting (Pruisken and Schäfer, 1982) and the interacting case (Finkel'stein, 1983a). The latter procedure requires some justification, which we shall come back to in the next subsection. Saddle-point solutions of the full model, Eqs. (3.92), have not been discussed until very recently (Belitz and Kirkpatrick, 1992). The most general saddle-point solution would be very complicated, as several different kinds of spontaneous symmetry breaking can occur. It is therefore expedient to restrict oneself initially to discussing different cases separately. We shall demonstrate the derivation of the noninteracting saddle-point solution and of disordered Hartree-Fock theory and BCS-Gorkov theory in the interacting case. Superconducting saddle-point solutions involving the particle-particle spin-triplet channel will be discussed in Sec. VIII. Magnetic solutions, as well as solutions describing interplay between magnetism and superconductivity, exist but have not been considered so far.

a. The saddle-point solution in the noninteracting case and Hartree-Fock theory

Let us consider the three terms Eq. (3.92a) contributes to the saddle-point equation. The first term is

$$\frac{\delta}{\delta^i_rQ^{\beta\alpha}_{mn}(\mathbf{x})}\int d\mathbf{y}\,\mathrm{tr}(Q(\mathbf{y}))^2=8\begin{bmatrix}+\\-\\-\\-\end{bmatrix}_r\begin{bmatrix}+\\-\\-\\-\end{bmatrix}_i{}^iQ^{\alpha\beta}_{nm}(\mathbf{x})\ , \tag{3.94}$$

with

$$\begin{bmatrix}+\\-\\-\\-\end{bmatrix}_\nu=\delta_{\nu0}-\sum_{\mu=1}^{3}\delta_{\nu\mu},\quad \nu=r,i\ .$$

In the second term, we denote the bare Green's function by

$$\mathcal{G}^0_{nm}{}^{\alpha\beta}(\mathbf{x},\mathbf{y})=-\delta_{\alpha\beta}\delta_{nm}(i\omega_n-H_0)^{-1}_{\mathbf{x},\mathbf{y}}$$
$$\equiv-\delta_{\alpha\beta}\delta_{nm}G^0_n(\mathbf{x}-\mathbf{y}) \tag{3.95a}$$

and expand the logarithm. This yields

$$\frac{\delta}{\delta^i_rQ^{\beta\alpha}_{mn}(\mathbf{x})}\int d\mathbf{y}\,\mathrm{tr}\ln\left[-(\mathcal{G}^0)^{-1}(\mathbf{y})\tau_0\otimes s_0+\frac{i}{2\tau}Q(\mathbf{y})\right]=-\frac{i}{2\tau}\mathcal{G}^0_{nm}{}^{\alpha\beta}(\mathbf{x},\mathbf{x})\mathrm{tr}\tau_r\mathrm{tr}s_i$$
$$-\left[\frac{i}{2\tau}\right]^2\sum_{\underline{r},\underline{i}}(\mathcal{G}^{0i}_{\underline{r}}Q\mathcal{G}^0)^{\alpha\beta}_{nm}(\mathbf{x},\mathbf{x})\mathrm{tr}\tau_r\tau_{\underline{r}}\mathrm{tr}s_is_{\underline{i}}$$
$$-\left[\frac{i}{2\tau}\right]^3\sum_{\underline{r},\underline{r}}\sum_{\underline{i},\underline{i}}\mathcal{G}^{0i}_{\underline{r}}Q\mathcal{G}^{0\underline{i}}_{\underline{r}}Q\mathcal{G}^0)^{\alpha\beta}_{nm}(\mathbf{x},\mathbf{x})\mathrm{tr}\tau_r\tau_{\underline{r}}\tau_{\underline{r}}\mathrm{tr}s_is_{\underline{i}}s_{\underline{i}}-\cdots\ . \tag{3.95b}$$

The right-hand side of Eq. (3.95b) has the structure of a Dyson equation. This identifies $iQ/2\tau$ as the self-energy, and

$$\mathcal{G}^{-1}=(\mathcal{G}^0)^{-1}(\tau_0\otimes s_0)-iQ/2\tau \tag{3.95c}$$

as the inverse dressed Green's function. This identification can be confirmed by retaining a source in the action, Eq. (3.31), while transforming to the Q variables (Pruisken and Schäfer, 1982). Finally, from the interaction term we have the contributions

$$\frac{\delta}{\delta {}^i_r Q^{\beta\alpha}_{mn}(\mathbf{x})}\int d\mathbf{y}[Q(\mathbf{y})\gamma^{(1)}Q(\mathbf{y})]=\delta_{\alpha\beta}\delta_{i0}\begin{bmatrix}+\\0\\0\\-\end{bmatrix}_r 32K^{(1)}\sum_{n_1n_2}\delta_{m+n_1,n+n_2}\,{}^i_r Q^{\alpha\alpha}_{n_1n_2}(\mathbf{x})\ , \tag{3.96a}$$

$$\frac{\delta}{\delta {}^i_r Q^{\beta\alpha}_{mn}(\mathbf{x})}\int d\mathbf{y}[Q(\mathbf{y})\gamma^{(2)}Q(\mathbf{y})]=-\delta_{\alpha\beta}\begin{bmatrix}0\\+\\+\\+\end{bmatrix}_i\begin{bmatrix}+\\0\\0\\-\end{bmatrix}_r 32K^{(2)}\sum_{n_1n_2}\delta_{m+n_1,n+n_2}\,{}^i_r Q^{\alpha\alpha}_{n_1n_2}(\mathbf{x})\ , \tag{3.96b}$$

$$\frac{\delta}{\delta {}^i_r Q^{\beta\alpha}_{mn}(\mathbf{x})}\int d\mathbf{y}[Q(\mathbf{y})\gamma^{(3)}Q(\mathbf{y})]=-\delta_{\alpha\beta}\delta_{i0}\begin{bmatrix}0\\+\\+\\0\end{bmatrix}_r 16K^{(3)}\sum_{n_1n_2}\delta_{m+n,n_1+n_2}\,{}^0_r Q^{\alpha\alpha}_{n_1n_2}(\mathbf{x})\ , \tag{3.96c}$$

$$\frac{\delta}{\delta {}^i_r Q^{\beta\alpha}_{mn}(\mathbf{x})}\int d\mathbf{y}[Q(\mathbf{y})\gamma^{(4)}Q(\mathbf{y})]=\delta_{\alpha\beta}\begin{bmatrix}0\\+\\+\\+\end{bmatrix}_i\begin{bmatrix}0\\+\\+\\0\end{bmatrix}_r 16\sum_{n_1n_2}K^{(4)}_{mn,n_1n_2}\delta_{m+n,n_1+n_2}\,{}^i_r Q^{\alpha\alpha}_{n_1n_2}(\mathbf{x})\ . \tag{3.96d}$$

In deriving Eq. (3.96d) we have made use of the symmetry relation, Eq. (3.53). Finally, $\delta S_{\rm int}$, Eq. (3.92i), gives a contribution proportional to

$$\frac{\delta}{\delta {}^i_r Q^{\beta\alpha}_{mn}(\mathbf{x})}\int d\mathbf{y}(\mathrm{tr}Q(\mathbf{y}))^2=\delta_{\alpha\beta}\delta_{i0}\delta_{r0}\delta_{nm}2\,\mathrm{tr}Q(\mathbf{x})\ . \tag{3.96e}$$

Inspection of Eqs. (3.94)–(3.96) shows that the first term on the right-hand side of Eq. (3.95b) is the only one that can provide an inhomogeneity (viz., the bare Green's function) in the equation for Q. If either $r\neq 0$ or $i\neq 0$, then there is no inhomogeneity, and the self-energy will be nonzero only if there is a broken symmetry, in which case Q serves as the order parameter. ${}^i_r Q\neq 0$, $r=0,3$, $i=1,2,3$ describes magnetism, a possibility that we shall ignore. ${}^0_r Q\neq 0$, $r=1,2$ describes spin-singlet superconductivity (see the next subsection), and ${}^i_r Q\neq 0$, $r=1,2$, $i=1,2,3$ describes spin-triplet superconductivity (see Sec. VIII). For now we restrict ourselves to the case with no broken symmetry and make the ansatz

$${}^i_r Q^{\alpha\beta}_{nm}(\mathbf{x})=\delta_{r0}\delta_{i0}\delta_{mn}\delta_{\alpha\beta}(-i2\tau)\Sigma_n\ . \tag{3.97}$$

This yields

$$\Sigma_n=\frac{-1}{2\pi N_F\tau}\sum_{\mathbf{k}}G_n(\mathbf{k})-\frac{8}{N_F}K^{(1)}T\tau\sum_n\Sigma_n\ , \tag{3.98a}$$

where we have omitted the contribution from Eq. (3.96e) since it vanishes in the replica limit. If, in addition to the restrictions inherent in the ansatz (3.97), we neglect an irrelevant real contribution to Σ_n that simply renormalizes the chemical potential, and consider the low-frequency limit and assume $\Sigma_n=\Lambda_n$ with $\Lambda_n=-\Lambda_{-n}$, we obtain

$$\Lambda_n=-\frac{-1}{2\pi N_F\tau}\sum_{\mathbf{k}}G_n(\mathbf{k})=\frac{i}{2\tau}\mathrm{sgn}(\omega_n)\ . \tag{3.98b}$$

This is the usual self-energy contribution from elastic impurity scattering, and Eq. (3.98b) in Eq. (3.97) gives the saddle-point solution of the noninteracting theory (Efetov *et al.*, 1980; Finkel'stein, 1983a). If we do not assume Σ_n to be an antisymmetric function of n, we find

$$\Sigma_n=\frac{-1}{2\pi N_F\tau}\sum_{\mathbf{k}}G_n(\mathbf{k})+\frac{4}{\pi N_F^2}K^{(1)}[1+O(T\tau)]\sum_{\mathbf{k}}T\sum_m G_m(\mathbf{k})\ . \tag{3.98c}$$

We remember that the particle-hole singlet interaction amplitude $K^{(1)}$ is the sum of an exchange and a direct contribution [this can be seen from Eqs. (3.92h) and (3.49)] and recognize Eq. (3.98c) as the Hartree-Fock equation in the presence of elastic scattering; see the diagrammatic representation in Fig. 13.

In accord with the discussion after Eq. (3.85) we have omitted corrections of $O(T\tau)$ in Eq. (3.98). As we have mentioned earlier, this is unnecessary. Let us go back to Eq. (3.85a), keep the correction term, and use Eq. (3.92h). This leads to the following replacement of $K^{(1)}$ in the action, Eq. (3.92d):

$$K^{(1)}\to K^{(1)}\left[1-\frac{8}{N_F}K^{(1)}T\tau\sum_n\right]^{-1}\ . \tag{3.99}$$

We make this substitution in Eq. (3.98a) and solve the integral equation for Σ_n. This yields

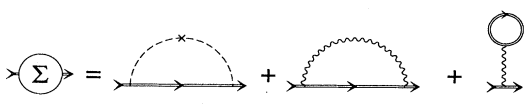

FIG. 13. Diagrammatic representation of the self-energy in the saddle-point approximation.

$$\Sigma_n=\frac{-1}{2\pi N_F\tau}\sum_{\mathbf{k}}G_n(\mathbf{k})+\frac{4}{\pi N_F^2}K^{(1)}\sum_{\mathbf{k}}T\sum_n G_n(\mathbf{k})\ , \tag{3.98c'}$$

which is Eq. (3.98c) with the correction terms *vanishing identically*. In the next subsection we shall find the analogous result for the particle-particle channel. We conclude that, at the level of the saddle-point solution, the terms proportional to $T\tau\Sigma_n$ drop out of the final results for observable quantities. Presumably this will also be true, order by order, in the loop expansion to be set up in the next section, but this has never been checked explicitly.

b. BCS-Gorkov theory as a saddle-point solution of the field theory

We now turn our attention to saddle-point solutions describing superconductivity. For simplicity, we neglect normal self-energies coming from the interaction by putting $K^{(1)}=K^{(2)}=0$. That is, we do not explicitly consider the Coulomb repulsion, and $K^{(3)}$ should be interpreted as containing the Coulomb pseudopotential: $K^{(3)}\propto-\lambda+\mu^*$. We also ignore the possibility of triplet superconductivity, $K^{(4)}=0$. We shall consider the case $K^{(4)}\neq0$ in Sec. VIII.

Since BCS-Gorkov theory (Abrikosov and Gorkov, 1958, 1959; Gorkov, 1959; see also Abrikosov *et al.*, 1975) is already fairly complicated, let us first derive BCS theory in the clean limit. Accordingly, we make the ansatz

$${}^{i}_{r}Q^{\beta\alpha}_{mn}(\mathbf{x})=\delta_{\alpha\beta}\delta_{r1}\delta_{i0}\delta_{m,-n}2\tau i\bar{\Delta}_m\ . \tag{3.100}$$

This ansatz implies that we work to leading order in $1/\tau$ only, as we have no disorder-related self-energy. The factor of i is convenient to make the gap function Δ_m real (notice that we have expanded in quaternions rather than in Pauli matrices, as is customary in superconductivity theory). Equation (3.100) in Eq. (3.93) yields

$$\bar{\Delta}_n=\frac{1}{2\pi N_F\tau}\sum_{\mathbf{k}}\tilde{F}_n^{+}(\mathbf{k})-\frac{4}{N_F}T\tau\tilde{K}^{(3)}\sum_m\bar{\Delta}_m\ . \tag{3.101a}$$

Here we have again kept the terms of $O(T\tau)$ in the action, and therefore replaced $K^{(3)}$ by [cf. Eq. (3.85b)]

$$\tilde{K}^{(3)}=K^{(3)}\left[1-\frac{4}{N_F}K^{(3)}T\tau\sum_n\right]^{-1}\ . \tag{3.101b}$$

The function $\tilde{F}^{+}$ is given by the infinite series

$$\tilde{F}_n^{+}(\mathbf{k})=\Delta_n G_n^0(\mathbf{k})G^0_{-n}(\mathbf{k})\times\sum_{\nu=0}^{\infty}(-)^{\nu}(\bar{\Delta}_n)^{2\nu}[G_n^0(\mathbf{k})G^0_{-n}(\mathbf{k})]^{\nu}\ . \tag{3.101c}$$

Here we have used $\bar{\Delta}_n=\bar{\Delta}_{-n}$, which follows from Eqs. (3.100) and (3.66c). We solve Eq. (3.101a) for $\Delta\equiv\Delta_n$ in the limit $\tau\to\infty$ and find

$$\Delta=-\frac{2}{\pi N_F^2}K^{(3)}\sum_{\mathbf{k}}T\sum_n\tilde{F}_n^{+}(\mathbf{k})\ . \tag{3.102a}$$

Equation (3.101c) for $\tilde{F}_n^{+}$ can be rewritten

$$\tilde{F}_n^{+}(\mathbf{k})=\Delta G^0_{-n}(\mathbf{k})\tilde{G}_n(\mathbf{k})\ , \tag{3.102b}$$

where

$$\tilde{G}_n(\mathbf{k})=G_n^0(\mathbf{k})-G_n^0(\mathbf{k})\Delta\tilde{F}_n^{+}(\mathbf{k})\ . \tag{3.102c}$$

Solving for $\tilde{F}_n^{+}$, one obtains the Gorkov function

$$\tilde{F}_n^{+}(\mathbf{k})=\frac{\Delta}{\omega_n^2+\xi_{\mathbf{k}}^2+\Delta^2}\ , \tag{3.103}$$

where $\xi_{\mathbf{k}}=\mathbf{k}^2/2m-\mu$, and the saddle-point equation takes the familiar form $\Delta=0$, or

$$1=\frac{2}{\pi N_F^2}|K^{(3)}|\sum_{\mathbf{k}}T\sum_n\frac{1}{\omega_n^2+\xi_{\mathbf{k}}^2+\Delta^2}\ . \tag{3.104}$$

Equation (3.104) is the gap equation of BCS theory (see, for example, Abrikosov *et al.*, 1975).

Now we repeat the calculation for arbitrary τ by including the disorder piece of the self-energy. Equation (3.100) gets replaced by

$${}^{i}_{r}Q^{\beta\alpha}_{mn}(\mathbf{x})=\delta_{\alpha\beta}\delta_{i0}2\tau i[-\delta_{nm}\delta_{r0}\Sigma_m+\delta_{m,-n}\delta_{r1}\bar{\Delta}_m]\ . \tag{3.105}$$

As a result of this more complicated ansatz, the equations for the Green's functions F^{+} and G will become more complicated, but the structure of the theory is the same as before, and the result will again be Eq. (3.104). Equation (3.105) in Eq. (3.93) yields

$$\bar{\Delta}_n=\frac{-2}{\pi N_F^2}K^{(3)}\sum_{\mathbf{k}}T\sum_n F_n^{+}(\mathbf{k})+\frac{1}{2\pi N_F\tau}\sum_{\mathbf{k}}F_n^{+}(\mathbf{k})\ , \tag{3.106a}$$

$$\Sigma_n=\frac{-1}{2\pi N_F\tau}\sum_{\mathbf{k}}G_n(\mathbf{k})\ . \tag{3.106b}$$

The functions F^{+} and G are most conveniently expressed in terms of $\tilde{F}^{+}$ and $\tilde{G}$ from clean BCS theory. The latter are given by Eqs. (3.102b) and (3.102c) and have to be augmented by the normal self-energy Σ in order to yield F^{+} and G. Terms linear in Σ are easily incorporated by writing (in symbolic notation)

$$F^{+}=\tilde{F}^{+}+\tilde{F}^{+}\Sigma\tilde{G}+\tilde{G}\Sigma\tilde{F}^{+}+O(\Sigma^2)\ , \tag{3.107a}$$

$$G=\tilde{G}+\tilde{G}\Sigma\tilde{G}-\tilde{F}^{+}\Sigma\tilde{F}^{+}+O(\Sigma^2)\ . \tag{3.107b}$$

For our purposes, $\tilde{F}=\tilde{F}^{+}$. This would change, for example, in the presence of a magnetic field. From the usual structure of Dyson's equation it is now obvious that the result to all orders in Σ can be obtained by dressing one of the BCS-Green's functions in the first-order terms with Σ. With frequency and wave-number dependence

put back in, we have

$$F_n^+(\mathbf{k})=\tilde{F}_n^+(\mathbf{k})+\tilde{F}_n^+(\mathbf{k})\Sigma_n G_n(\mathbf{k})+\tilde{G}_{-n}(\mathbf{k})\Sigma_{-n}F_n^+(\mathbf{k}) , \tag{3.108a}$$

$$G_n(\mathbf{k})=\tilde{G}_n(\mathbf{k})+\tilde{G}_n(\mathbf{k})\Sigma_n G_n(\mathbf{k})-\tilde{F}_n(\mathbf{k})\Sigma_{-n}F_n^+(\mathbf{k}) . \tag{3.108b}$$

A graphic representation of Eqs. (3.108) is given in Fig. 14. These are the Gorkov equations in standard form, and the remainder of the discussion can be found in textbooks (e.g., Abrikosov *et al.*, 1975). The final result is

$$F_n^+(\mathbf{k})=\frac{1}{\eta_n}\frac{\Delta}{\omega_n^2+(\xi_\mathbf{k}/\eta_n)^2+\Delta^2} , \tag{3.109a}$$

$$G_n(\mathbf{k})=\frac{-1}{\eta_n}\frac{i\omega_n+\xi_k/\eta_n}{\omega_n^2+(\xi_\mathbf{k}/\eta_n)^2+\Delta^2} , \tag{3.109b}$$

with

$$\eta_n=1+\frac{1/\Delta}{2\pi N_F\tau}\sum_\mathbf{k}F_n^+(\mathbf{k}) , \tag{3.109c}$$

and the gap Δ still determined by Eq. (3.104).

A remarkable aspect of this result is that the gap is independent of the disorder and therefore so are all thermodynamic properties. This fact is often referred to as Anderson's theorem (Anderson, 1959b). The physical reason behind Anderson's theorem, namely, particle number conservation, is completely obscured in the above derivation. It can be seen more clearly by expressing the superconducting T_c, and hence Δ, in terms of a density-density correlation function for noninteracting electrons (see, for example, Belitz, 1990, and Sec. VII below). In the absence of magnetic fields or magnetic impurities, the correlation function is to be taken in the homogeneous limit, $q\to 0$, where, due to particle number conservation, it depends on disorder only through the DOS; see Eq. (2.36).

3. The generalized nonlinear sigma model

a. An effective theory for diffusion modes

Equations (3.92) have been derived for slowly varying matrices Q. It is therefore natural to expand in powers of the gradient of Q and in powers of the external frequency. Furthermore, the second term on the right-hand side of Eq. (3.92a) contains a term linear in Q which can be eliminated by a suitable shift of Q. It is possible, and physically appealing, to expand about the Hartree-Fock saddle-point solution, Eqs. (3.97) and (3.98c). We notice two points, however. (1) With our static potential, the Hartree-Fock saddle point, Eq. (3.98c), differs from the noninteracting saddle point, Eq. (3.98b), only by an additive constant. The interaction therefore just shifts the chemical potential and is of no deep significance. (2) Keeping this shift, and the resulting self-consistent change of the first term in Eq. (3.98c), amounts to an approximate consideration of Fermi-liquid corrections to the parameters of the effective theory to be derived. Since the effective theory should be interpreted as containing the *exact* Fermi-liquid corrections stemming from the nonhydrodynamic regime (see Sec. III.B.3.d below), it makes little sense to include these approximate effects explicitly.

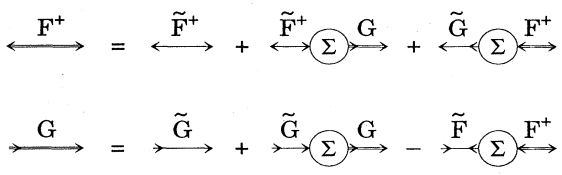

FIG. 14. Diagrammatic representation of the Gorkov equations.

With these points in mind, we write (Finkel'stein, 1983a)

$$Q=Q_{\rm SP}+\tilde{Q}-2\tau\Omega , \tag{3.110a}$$

where

$$^{i}_{r}(Q_{\rm SP})^{\alpha\beta}_{nm}(\mathbf{x})=\delta_{r0}\delta_{i0}\delta_{nm}\delta_{\alpha\beta}{\rm sgn}(\omega_n) \tag{3.110b}$$

and

$$\Omega^{\alpha\beta}_{nm}=\delta_{\alpha\beta}\delta_{nm}\omega_n\tau_0\otimes s_0 , \tag{3.110c}$$

and expand the action in a Taylor series about $Q_{\rm SP}$. The terms linear in $Q_{\rm SP}$ vanish, the first term on the right-hand side of Eq. (3.92a) yields a term proportional to ${\rm tr}(\tilde{Q}^2)$ and one proportional to ${\rm tr}(\Omega\tilde{Q})$, and the only nontrivial contribution comes from the tr ln term in Eq. (3.92a). To second order in $\tilde{Q}$ one finds (McKane and Stone, 1981; Finkel'stein, 1983a)

$$\begin{aligned}{\rm tr}\ln[i\omega-H_0+iQ(\mathbf{x})/2\tau]&={\rm tr}\ln[-H_0+iQ_{\rm SP}(\mathbf{x})/2\tau]\\&\quad+\frac{1}{8\tau^2}{\rm tr}[\mathcal{G}_{\rm SP}(\mathbf{x})\tilde{Q}(\mathbf{x})\mathcal{G}_{\rm SP}(\mathbf{x})\tilde{Q}(\mathbf{x})]\\&\quad+\pi N_F{\rm tr}[\Omega\tilde{Q}(\mathbf{x})] ,\end{aligned} \tag{3.111a}$$

where $\mathcal{G}_{\rm SP}$ is the saddle-point Green's function,

$$\mathcal{G}_{\rm SP}=[(\mathcal{G}^0)^{-1}(\tau_0\otimes s_0)-iQ_{\rm SP}/2\tau]^{-1} , \tag{3.111b}$$

in the limit of zero frequency. Adding the remaining terms in Eq. (3.92a) we find

$$\begin{aligned}S[\tilde{Q}]\simeq&\frac{-\pi N_F}{8\tau}\sum_\mathbf{k}{\rm tr}[\tilde{Q}(\mathbf{k})\tilde{Q}(-\mathbf{k})]\\&\times\left[1-\frac{1}{2\pi N_F\tau}\sum_\mathbf{p}G_{\rm SP}(\mathbf{p})G_{\rm SP}(\mathbf{p}-\mathbf{k})\right]\\&+\frac{\pi N_F}{2}\int d\mathbf{x}\,{\rm tr}[\Omega\tilde{Q}(\mathbf{x})]+S_{\rm int}[\tilde{Q}-2\tau\Omega] .\end{aligned} \tag{3.112}$$

Here G_{SP} is related to $\mathcal{G}_{SP}$ in the same way as G^0 to $\mathcal{G}^0$ in Eq. (3.95a). For reasons that will become clear below, we first restrict ourselves to fluctuations $\tilde{Q}_{nm}$ with $n \geq 0$ and $m \leq -1$ or $n \leq -1$ and $m \geq 0$. For these cases the term in the square brackets in Eq. (3.112) is readily identified as an essential ingredient in the diagrammatic derivation of the diffusion pole in Sec. II. We have [cf. Eqs. (2.31)–(2.32)]

$$1-\frac{1}{2\pi N_F \tau}\sum_{\mathbf{p}} G_{SP}(\mathbf{p})G_{SP}(\mathbf{p}-\mathbf{k}) = -1-I_0(\mathbf{k},\Omega=0) = \tau D k^2 + O(k^4) , \tag{3.113}$$

where $D=v_F^2\tau/d$ is the bare diffusion constant in d dimensions. Neglecting terms of order $(\nabla\tilde{Q})^4$, Ω^2, and $\Omega(\nabla\tilde{Q})^2$, we now have

$$S[\tilde{Q}] = -\frac{\pi}{8}N_F D \int d\mathbf{x}\, \mathrm{tr}(\vec{\nabla}\tilde{Q}(\mathbf{x}))^2 + \frac{\pi N_F}{2}\int d\mathbf{x}\, \mathrm{tr}[\Omega\tilde{Q}(\mathbf{x})] + S_{\mathrm{int}}[\tilde{Q}] . \tag{3.114}$$

So far we have neglected Ω in the argument of S_{int}. We shall see in the next subsection that $S_{\mathrm{int}}[\tilde{Q}]$ is effectively proportional to Ω, so the neglected terms are of $O(\Omega^2)$. We have also ignored fluctuations $\tilde{Q}_{nm}$ with $n,m \geq 0$ or $n,m \leq -1$. It is easily shown that, in the Gaussian approximation, fluctuations Q_{nm} with $n,m \geq 0$ or $n,m \leq -1$ are massive in the sense that they are not singular in the long-wavelength, small-frequency limit. However, these fluctuations cannot be simply neglected because at higher order the massive Q's couple to the massless ones. Technically the massive Q's need to be integrated out. This can be most simply accomplished by putting two constraints on the matrix $\tilde{Q}=Q-Q_{SP}+2\tau\Omega$. To this end we note that, for zero external frequency, $\Omega=0$, which implies $S_{\mathrm{int}}=0$, homogeneous fluctuations of Q that satisfy $Q^2=(Q_{SP})^2=1$ do not change the free energy. In this case, only the gradients of Q will contribute. This implies that the first term in Eq. (3.114) is adequate if we restrict ourselves to matrices that obey

$$Q^2 = 1 , \tag{3.115}$$

where 1 denotes the unit matrix. Furthermore, it is clear from Eq. (3.92i) that fluctuations with $\mathrm{tr}Q \neq \mathrm{tr}Q_{SP}=0$ greatly increase the energy. We note that this is true also in the absence of interactions, as can be seen from Eq. (3.76a). One can suppress these massive fluctuations by requiring

$$\mathrm{tr}Q = 0 . \tag{3.116}$$

The net result is that the $\tilde{Q}$ matrix in Eq. (3.114) can be regarded as a general Q_{nm} matrix if the constraints given by Eqs. (3.115) and (3.116) are enforced.

Equations (3.114)–(3.116) (with $S_{\mathrm{int}}=0$) were originally written down by Wegner (1979). They constitute a nonlinear sigma model, i.e., a free-field theory with a constraint of the form of Eq. (3.115). Models of this form had originally been introduced in the theory of beta decay (Gell-Mann and Lévy, 1960). They were then applied to condensed-matter physics as models for ferromagnetism (Polyakov, 1975; Brézin and Zinn-Justin, 1976; Nelson and Pelcovits, 1977). In these models the analog of Q was an n-component vector. Wegner argued phenomenologically, drawing analogies to ferromagnetism, that the matrix nonlinear sigma model should describe the soft modes of disordered electron systems. His reasoning was corroborated by technical derivations of the model (Schäfer and Wegner, 1980; McKane and Stone, 1981; Pruisken and Schäfer, 1982). These authors derived the model by explicitly integrating out the massive modes, making heavy use of the symmetry inherent in the model and of the way it is broken by the external frequency. The interaction term and its consequences were not considered at this technical level. The model with the interaction term included should therefore be viewed as somewhat more phenomenological than the pure localization model. This also shows in the renormalization-group analysis, which we shall discuss below.

Equations (3.114)–(3.116) are sufficient to completely parametrize the model. The constraints (3.115), (3.116), and the Hermiticity requirement (3.63) can be eliminated by parametrizing the matrix $\tilde{Q}$ (Grilli and Sorella, 1988; Belitz and Kirkpatrick, 1989a) by

$$\tilde{Q} = \begin{array}{c|cc|c} & m\geq 0 & m<0 & \\ \hline & (1-qq^+)^{1/2}-1 & q & n\geq 0 \\ & q^+ & -(1-q^+q)^{1/2}+1 & n<0 \end{array} \tag{3.117a}$$

Here the q are matrices with spin-quaternion-valued elements $q_{nm}^{\alpha\beta}$; $n=0,1,\ldots$; $m=-1,-2,\ldots$. Like the matrix Q, they are conveniently expanded in the $\tau\otimes s$ basis,

$$q_{nm}^{\alpha\beta} = \sum_{r=0}^{3}\sum_{i=0}^{3} {}^{i}_{r}q_{nm}^{\alpha\beta}\tau_r\otimes s_i . \tag{3.117b}$$

We note that the q do *not* obey the symmetry relations (3.66) for the Q.

The parametrization of $\tilde{Q}$ is not unique. A different possibility, which has been used widely (Polyakov, 1975; Efetov *et al.*, 1980; Finkel'stein, 1983a), is

$$\tilde{Q} = Q_{SP}[\exp W - 1] , \tag{3.117c}$$

with

$$W = \begin{array}{c|cc|c} & m\geq 0 & m<0 & \\ \hline & 0 & w & n\geq 0 \\ & -w^+ & 0 & n<0 , \end{array} \tag{3.117d}$$

and the matrix elements of the w again spin quaternions. While the two parametrizations are equivalent, Eq. (3.117a) has certain technical advantages, as we shall see.

In terms of $\tilde{Q}$ the action reads

$$S[\tilde{Q}]=\frac{-1}{2G}\int d\mathbf{x}\,\mathrm{tr}(\vec{\nabla}\tilde{Q}(\mathbf{x}))^2 +2H\int d\mathbf{x}\,\mathrm{tr}(\Omega\tilde{Q}(\mathbf{x}))+S_{\mathrm{int}}[\tilde{Q}]\ . \tag{3.118}$$

The coupling constants are $G=4/\pi N_F D=8/\pi\sigma$, with σ the bare conductivity, and $H=\pi N_F/4$, which plays the role of a frequency coupling parameter. S_{int} is given by Eqs. (3.92c)–(3.92g). Notice that δS_{int} in Eq. (3.92c) has been dropped from the effective action.

b. Gaussian theory

With the help of Eq. (3.117a) one can expand the action in powers of q,

$$S[\tilde{Q}]=\sum_{n=2}^{\infty}S_n[q]\ , \tag{3.119}$$

where $S_n[q]\propto q^n$. We first concentrate on the Gaussian part of the action:

$$S_2[q]=\frac{-4}{G}\int_{\mathrm{p}}\sum_{r,i}\sum_{\substack{1,2\\3,4}}{}^{i}_{r}q_{12}(\mathbf{p})\,{}^{i}_{r}M_{12,34}(p)\,{}^{i}_{r}q_{34}(-\mathbf{p})\ . \tag{3.120a}$$

Here $\int_{\mathrm{p}}\equiv\int d\mathbf{p}/(2\pi)^d$, and $1=(n_1,\alpha_1)$, etc. The matrix M is given by

$${}^{i}_{0,3}M_{12,34}(p)=\delta_{1-2,3-4}\{\delta_{13}(p^2+GH(\omega_{n_1}-\omega_{n_2}))+\delta_{\alpha_1\alpha_2}\delta_{\alpha_1\alpha_3}2\pi TGK^{(\nu_i)}\}\ , \tag{3.120b}$$

where $\nu_0=s$, $\nu_{1,2,3}=t$, and

$${}^{i}_{1,2}M_{12,34}(p)=-\delta_{1+2,3+4}\{\delta_{13}(p^2+GH(\omega_{n_1}-\omega_{n_2}))+\delta_{\alpha_1\alpha_2}\delta_{\alpha_1\alpha_3}2\pi TGK^{c(\nu_i)}_{n_1n_2,n_3n_4}\}\ . \tag{3.120c}$$

In Eq. (3.120c) we have included the frequency dependence of the interaction coupling parameter since, as mentioned above, $K^{c(t)}\equiv 0$ if it is assumed to be frequency independent. In Sec. VI we shall see that, even if one starts with $K^{c(t)}=0$, the renormalization group *generates* a frequency-dependent, nonzero $K^{c(t)}$. The consequences of this have been analyzed at the level of the saddle-point solution (see Sec. VIII), but not yet at the level of the loop expansion. In the remainder of this section, and in the following three sections, we shall therefore assume $K^{c(t)}=0$.

The inverse of the matrix M determines the Gaussian propagators. Since M is diagonal in charge, spin, and replica labels it is sufficient to consider the frequency dependence. Consider operators with structures

$${}^{0}P_{n_1n_2,n_3n_4}=a(n_1-n_2)\delta_{n_1n_3}\delta_{n_2n_4}+b\delta_{n_1-n_2,n_3-n_4}\ , \tag{3.121a}$$

$${}^{1}P_{n_1n_2,n_3n_4}=a(n_1-n_2)\delta_{n_1n_3}\delta_{n_2n_4}+b\delta_{n_1+n_2,n_3+n_4}\ , \tag{3.121b}$$

with a and b constants. It is readily checked that the inverses of these operators are

$$({}^{0}P^{-1})_{n_1n_2,n_3n_4}=\frac{1}{a(n_1-n_2)}\delta_{n_1n_3}\delta_{n_2n_4}+\delta_{n_1-n_2,n_3-n_4}\frac{1}{n_1-n_2}\left[\frac{1}{(a+b)(n_1-n_2)}-\frac{1}{a(n_1-n_2)}\right]\ , \tag{3.121c}$$

$$({}^{1}P^{-1})_{n_1n_2,n_3n_4}=\frac{1}{a(n_1-n_2)}\delta_{n_1n_3}\delta_{n_2n_4}-\delta_{n_1+n_2,n_3+n_4}\frac{b}{a(n_1-n_2)a(n_3-n_4)}\frac{1}{1+f(n_1+n_2)}\ , \tag{3.121d}$$

where

$$f(n)=b\sum_{n_1,n_2}\frac{1}{a(n_1-n_2)}\delta_{n_1+n_2,n}\ . \tag{3.121e}$$

We note that with constant a, b, and unrestricted frequency sums, the formal expression for $f(n)$ is logarithmically divergent in the ultraviolet (UV) for all $\mathbf{p}$ and n and in the infrared (IR) for $\mathbf{p},n\to 0$. The UV divergence is due to the unphysical treatment of the high-frequency behavior and should be cut off by a large cutoff N, while the low-frequency singularity is the usual divergence one obtains in the particle-particle channel: for $b<0$ there is a Cooper instability in the denominator of the last term of Eq. (3.121d). These singularities are independent of the dimensionality and of any loop expansion. For $b>0$ it is, naively, tempting to let $f(n)\to\infty$ at the end of the calculations and thereby neglect terms that are logarithmically small. However, Kirkpatrick and Belitz (1993) have recently argued that these terms give rise to logarithmic corrections to scaling and to some asymptotic power laws near the metal-insulator transition whenever the particle-particle channel is present. We shall discuss this in Secs. V and VI. For now, we carry the formal expression for $f(n)$ through the calculations and deal with the fact that the $f(n)$ is logarithmically large at the end.

With the help of Eqs. (3.121), we can immediately write down the inverse of the matrix M, which determines the

Gaussian propagators,

$$\langle {}^{i}_{r}q_{12}(\mathbf{p}_1){}^{j}_{s}q_{34}(\mathbf{p}_2)\rangle^{(2)}=\int D[q]\,{}^{i}_{r}q_{12}(\mathbf{p}_1){}^{j}_{s}q_{34}(\mathbf{p}_2)e^{S_2[q]}\Big/\int D[q]e^{S_2[q]}$$
$$=\frac{G}{8}\delta_{rs}\delta_{ij}\delta(\mathbf{p}_1+\mathbf{p}_2){}^{i}_{r}M^{-1}_{12,34}(p_1)\ , \tag{3.122a}$$

with

$${}^{i}_{0,3}M^{-1}_{12,34}(\mathbf{p})=\delta_{1-2,3-4}\left\{\delta_{13}\mathcal{D}_{n_1-n_2}(\mathbf{p})+\frac{\delta_{\alpha_1\alpha_2}}{n_1-n_2}\Delta\mathcal{D}^{\nu_i}_{n_1-n_2}(\mathbf{p})\right\}\ , \tag{3.122b}$$

$${}^{i}_{1,2}M^{-1}_{12,34}(\mathbf{p})=-\delta_{1+2,3+4}\left\{\delta_{13}\mathcal{D}_{n_1-n_2}(\mathbf{p})-\delta_{i0}\mathcal{D}_{n_1-n_2}(\mathbf{p})\mathcal{D}_{n_3-n_4}(\mathbf{p})\frac{\delta_{\alpha_1\alpha_2}G2\pi TK^{c(s)}}{1+G2\pi TK^{c(s)}f_{n_1+n_2}(\mathbf{p})}\right\}\ , \tag{3.122c}$$

where $\nu_0=s$, $\nu_{1,2,3}=t$ and

$$f_n(\mathbf{p})=\sum_{n_1 n_2}\delta_{n,n_1+n_2}\mathcal{D}_{n_1-n_2}(\mathbf{p})\ . \tag{3.122d}$$

Here we have introduced the propagators

$$\mathcal{D}_n(\mathbf{p})=[\mathbf{p}^2+GH\Omega_n]^{-1}\ , \tag{3.122e}$$

$$\mathcal{D}^{s,t}_n(\mathbf{p})=[\mathbf{p}^2+G(H+K^{s,t})\Omega_n]^{-1}\ , \tag{3.122f}$$

$$\Delta\mathcal{D}^{s,t}_n(\mathbf{p})=\mathcal{D}^{s,t}_n(\mathbf{p})-\mathcal{D}_n(\mathbf{p})\ . \tag{3.122g}$$

A physical interpretation of Eqs. (3.122e) and (3.122f) will be given in Sec. III.B.3.d below. Equations (3.122b) and (3.122c) can be put into a more standard form by summing over n_3 and n_4,

$$\sum_{n_3,n_4}{}^{i}_{0,3}M^{-1}_{12,34}(\mathbf{p})=(1-\delta_{\alpha_1\alpha_2})\mathcal{D}_{n_1-n_2}(\mathbf{p})+\delta_{\alpha_1\alpha_2}\mathcal{D}^{\nu_i}_{n_1-n_2}(\mathbf{p})\ , \tag{3.123a}$$

$$\sum_{n_3,n_4}{}^{i}_{1,2}M^{-1}_{12,34}(\mathbf{p})=-(1-\delta_{\alpha_1\alpha_2})\mathcal{D}_{n_1-n_2}(\mathbf{p})-\delta_{\alpha_1\alpha_2}(1-\delta_{i0})\mathcal{D}_{n_1-n_2}(\mathbf{p})-\frac{\delta_{i0}\delta_{\alpha_1\alpha_2}\mathcal{D}_{n_1-n_2}(\mathbf{p})}{1+G2\pi TK^{c(s)}f_{n_1+n_2}(\mathbf{p})}\ . \tag{3.123b}$$

Examining the various terms in Eqs. (3.123), we see that all of them have a standard propagator structure except for the last contribution in Eq. (3.123b). We also note that this term is logarithmically small compared to the others.

c. Loop expansion and the renormalization group

The general scheme of the field-theoretic treatment of the problem is now as follows. Let us consider a q-q correlation function, which we write symbolically as

$$G^{(2)}=\langle qq\rangle=\frac{1}{Z}\int D[q]qqe^{\sum_{n=2}^{\infty}S_n[q]}=\frac{1}{Z}\int D[q]qqe^{S_2[q]}\left[1+S_4[q]+\frac{1}{2}(S_3[q])^2+S_6[q]+\cdots\right]\ . \tag{3.124}$$

All of what follows applies to correlations of more than two q's as well. Equation (3.124) expresses $G^{(2)}$ in terms of an infinite series of Gaussian correlation functions which, by means of Wick's theorem (e.g., Abrikosov *et al.*, 1975) and Eqs. (3.122), are easily calculated. It is convenient to depict the terms in this series graphically. Draw $S_n[q]$ as a point from which n lines emerge, with Gaussian propagators as joined lines. Then the four terms shown in Eq. (3.124) are represented graphically in Fig. 15. The expansion for $G^{(2)}$ contains both connected and disconnected diagrams, as in Fig. 15. However, according to a general theorem, $G^{(2)}$ is easily reconstructed from the connected propagator $G^{(2)}_c$, which is defined as the sum of all connected diagrams that contribute to $G^{(2)}$ (see, for example, Le Bellac, 1991). The number of diagrams to be considered can be further reduced by studying the two-point vertex function $\Gamma^{(2)}$, defined as the inverse (in a matrix sense) of $G^{(2)}_c$. From Dyson's equation we have

$$\Gamma^{(2)}=\Gamma^{(2)}_G-\Sigma^{(2)}\ , \tag{3.125}$$

with $\Gamma^{(2)}_G$, the Gaussian vertex, the inverse of the Gauss-

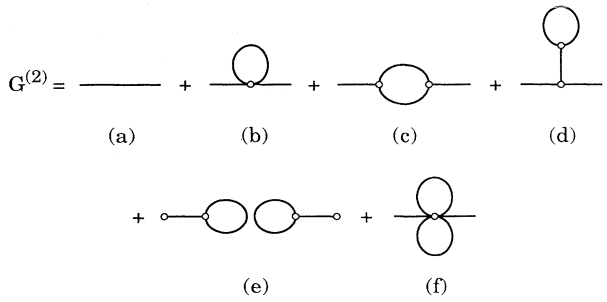

FIG. 15. Diagrams up to two-loop order for the propagator $G^{(2)}$.

ian propagator, and $\Sigma^{(2)}$ the mass operator (or self-energy). $\Sigma^{(2)}$ is one-particle irreducible, i.e., it contains no diagrams that can be cut into two disconnected parts by cutting only one line. Of the six diagrams in Fig. 15, only (a), (b), (c), and (f) contribute to $\Gamma^{(2)}$.

The diagrams can be classified with respect to the number of closed loops they contain. Using the Gaussian propagators, and counting powers of the coupling constant G, one easily sees that the number of loops is correlated with the number of powers of G. Equation (3.124) therefore provides a systematic expansion in powers of G, which is known as the loop expansion. It provides a perturbation expansion for the propagators or the vertices and therewith for the coupling constants of the theory. This expansion can be reproduced term by term by means of many-body perturbation theory, although in the framework of the latter it is much harder to find all diagrams contributing to a given order. In the limit $d \to 2$, the loop expansion will contain divergences of the type we already discussed in Sec. II. These divergences signal a breakdown of perturbation theory and the need for a resummation. Proper handling of this requires the use of the renormalization group (RG).

This is not the place to review the RG. Readers unfamiliar with this technique are directed to the many reviews and books on the subject (e.g., Wilson and Kogut, 1974; Amit, 1984; Itzykson and Droeffe, 1989; Zinn-Justin, 1989; Le Bellac, 1991). Work on the nonlinear sigma model has used both the momentum-shell version (Wilson and Kogut, 1974) of the RG and the physically less intuitive, but more powerful, field-theoretic version (e.g., Zinn-Justin, 1989). Examples of the former are Polyakov (1975), Nelson and Pelcovits (1977), and Efetov *et al.* (1980). Examples of the latter are Brézin *et al.* (1976) and Amit *et al.* (1978). Since the momentum-shell method is impractical at higher than one-loop order, and in Secs. V–VII we shall deal with higher orders in the loop expansion, we shall use the field-theoretic method in this review.

We note at this point that the renormalizability of the model, Eq. (3.118), in the presence of interactions has so far not been proven. The model with $S_{\rm int}=0$ is known to be renormalizable with two renormalization constants, one for G and one field renormalization (Brézin *et al.*, 1976). All calculations performed to date support the hypothesis that the full model is renormalizable with one additional renormalization constant for H and one for each interaction amplitude. In the remainder of this review we shall assume that this is true. We shall further discuss the open problem of renormalizability in Sec. X.

d. Fermi-liquid corrections and identification of observables

So far our derivation of the effective action, Eq. (3.118), has taken into account the hydrodynamic region only. Finkel'stein (1983a) has suggested a way to improve on this.

The nonhydrodynamic region leads to Fermi-liquid corrections to the parameters of the effective theory. Within Fermi-liquid theory (e.g., Negele and Orland, 1988; Pines and Nozières, 1989) the charge- and spin-polarization clouds that a particle creates around itself by charge- and spin-polarizing its environment are considered together with the particle as a quasiparticle (Landau, 1958). The quasiparticles have kinetic and thermodynamic properties that differ from those of the bare particle and that are parametrized by means of the Fermi-liquid parameters F_n^s; $n=0,1\ldots$ and F_n^a; $n=0,1,\ldots$. The isothermal density susceptibility, spin susceptibility, and specific heat of the quasiparticles are given by

$$\partial n/\partial\mu=(k_F m^*/\pi^2)/(1+F_0^s)\ , \tag{3.126a}$$

$$\chi_s=(\mu_B^2 k_F m^*/\pi^2)/(1+F_0^a)\ , \tag{3.126b}$$

$$C_V=(k_F m^*/3)T\ . \tag{3.126c}$$

Here μ_B is the Bohr magneton, and m^* is the effective mass,

$$m^*=m(1+F_1^s/3)\ . \tag{3.126d}$$

N_F and D in Eq. (3.114) should then be interpreted as ($d=2,3$)

$$D=v_F^2\tau/d\ , \tag{3.127a}$$

$$N_F=\frac{m^*}{2\pi}(k_F/\pi)^{d-2}\ , \tag{3.127b}$$

with $v_F=k_F/m^*$. Similarly, the bare interaction amplitudes in $S_{\rm int}$ should be interpreted as containing the appropriate Fermi-liquid corrections.

We now turn to the interpretation of the coupling constants of the field theory in terms of observable quantities. Let us first reconsider the Gaussian theory, Sec. III.B.3.b above. The three Gaussian propagators $\mathcal{D}$, $\mathcal{D}^s$, and $\mathcal{D}^t$, Eqs. (3.122e), (3.122f), and (3.122g), have the typical diffusion pole structure discussed in Sec. II, and we can identify the quantities multiplying the frequency as inverse diffusion constants. At the Gaussian level, they are determined by the bare coupling constants. $\mathcal{D}^s$ and $\mathcal{D}^t$ obviously describe charge and spin diffusion, and the respective bare diffusion constants are

$$D_c=1/G(H+K_s)\ , \tag{3.128a}$$

$$D_s = 1/G(H+K_t) \ . \tag{3.128b}$$

While this identification so far holds only at the Gaussian level, it is plausible that it should be exact. The reason is that the diffusion poles are a direct consequence of the conservation laws, as we have seen in Sec. II. In any theory that respects the conservation of charge, spin, and energy, the appropriate inverse propagators must be linear in ω_n and $\mathbf{p}^2$ for small $|\mathbf{p}|$ and ω_n, and the coefficient of the $\mathbf{p}^2$ term will be the diffusion coefficient.[8] Indeed, Castellani, DiCastro, Lee, Ma, Sorella, and Tabet (1986) have used the charge and spin conservation laws (in the form of Ward identities) to prove that Eqs. (3.128a) and (3.128b) hold exactly. This ensures that the same combinations of renormalized coupling constants will yield the renormalized diffusion constants. The interpretation of $\mathcal{D}$ is less clear, even at the Gaussian level. The fact that energy is the only remaining conserved quantity, and that $\mathcal{D}$ represents a third diffusion pole, independent of $\mathcal{D}^s$ and D^t, leads one to suspect that GH can be identified with the inverse heat diffusion coefficient. Castellani *et al.* (1987, 1988) have invoked the corresponding Ward identity for heat transport to show that $\mathcal{D}$ indeed contains the heat diffusion constant. We therefore have for the heat diffusion constant

$$D_h = D = 1/GH \ . \tag{3.128c}$$

Again, this not only holds at the Gaussian level, but is an exact identity.

From the Einstein relation, Eq. (2.21), we know that the charge conductivity and the charge diffusion constant are related via $\sigma_c = D_c(\partial n/\partial\mu)$. If we combine this with $G = 8/\pi\sigma_c$, we find that $H+K_s$ is proportional to the isothermal density susceptibility,

$$H+K_s = \frac{\pi}{8}(\partial n/\partial\mu) \ . \tag{3.129a}$$

The relation (3.129a) is important because the compressibility $\partial n/\partial\mu$ is not expected to be critically affected by disorder. Physically, one expects n to be a smooth function of μ even at a metal-insulator transition. Mathematically, $\partial n/\partial\mu$ does not acquire diffusion corrections in perturbation theory (Finkel'stein, 1983a; Sota and Suzuki, 1989). One therefore expects, though this has never been rigorously proven, that $H+K_s$ will not be renormalized.

In a completely analogous fashion, $H+K_t$ and H determine the spin and thermal susceptibilities, respectively. Finkel'stein (1984b), and Castellani, DiCastro, Lee, Ma, Sorella, and Tabet (1984, 1986), have shown that

$$H+K_t = (\pi/4\mu_B^2)\chi_s \ . \tag{3.129b}$$

Castellani and DiCastro (1986), and Castellani *et al.* (1987, 1988), have also established the relation between H and the specific-heat coefficient,

$$H = \frac{3}{4\pi}C_V/T \ , \tag{3.129c}$$

with C_V the specific heat at constant volume. Equations (3.129), like Eqs. (3.128), are exact identities that hold order by order in perturbation theory. Together, they provide an interpretation of all of the coupling constants of the field theory except for $K^{c(s)}$. It does not seem possible to relate this interaction amplitude to an observable quantity.

Let us compare the exact expressions for the susceptibilities, Eqs. (3.129), with their counterparts within Landau theory, Eqs. (3.126). If we substitute the Landau expressions for $\partial n/\partial\mu$, χ_s, and C_V in Eqs. (3.129) we obtain

$$H = \frac{\pi N_F}{4}(1+F_1^s/3) \ , \tag{3.130a}$$

$$K_s/H = -F_0^s/(1+F_0^s) = -A_0^s \ , \tag{3.130b}$$

$$K_t/H = -F_0^a/(1+F_0^a) = -A_0^a \ , \tag{3.130c}$$

with $A_0^{s,a}$ the Landau scattering amplitudes (see Abrikosov *et al.*, 1975). Notice that Eq. (3.130a) is consistent with the bare value of H, Eq. (3.118), and the reinterpretation of N_F in that expression, Eq. (3.127b). We also note that F_0^s and K_t are positive, while F_0^a and K_s are negative. Equations (3.130) suggest that H, K_s, and K_t are generalizations of the Fermi-liquid parameters F_1^s, A_0^s, and A_0^a. This idea has been developed and formalized by Castellani and DiCastro (1985, 1986), Castellani, Kotliar, and Lee (1987), and Castellani *et al.* (1987, 1988). These authors have shown that one can indeed construct a consistent Landau theory for quasiparticles in a disordered system that makes the analogy precise. In this picture, the quasiparticle diffusion constant is given by D_h, the quasiparticle DOS by $2H/\pi$, and the quasiparticle weight by $\pi N(0)/2H$ with $N(0)$ the tunneling DOS. Finally, the quasiparticle lifetime is given by $\tau_{\rm ph} 2H/\pi N_F$, where $\tau_{\rm ph}$ is the dephasing time discussed by Castellani, DiCastro, Kotliar, and Lee (1986).

Apart from the transport coefficients, we shall also be interested in the tunneling DOS. This is given by the expectation value of the diagonal elements of Q (Finkel'stein, 1983a),

$$N(\Omega) = N_F \mathrm{Re}\langle {}^0_0 Q^{\alpha\alpha}_{nn}(\mathbf{x})\rangle|_{i\Omega_n \to \Omega + i0} \ . \tag{3.131}$$

Here Ω represents a frequency (or energy) measured from the Fermi surface, and the average is to be taken with the full action. Note that the tunneling DOS is given by the one-point propagator, while the transport coefficients are determined by the two-point propagators.

[8] We note that the propagators $\mathcal{D}^s$, $\mathcal{D}^t$, and $\mathcal{D}$ are *not* the physical charge, spin, and heat density correlation functions. One can see this by calculating the latter by adding an appropriate source term to the action. Therefore the plausibility argument given above, strictly speaking, also assumes renormalizability of the theory with one renormalization constant per coupling constant. See also Sec. III.B.3.c above.

e. Long-range interaction

We now turn to the problem of the long-range nature of the Coulomb interaction, which we have ignored so far. A Coulomb potential

$$v(\mathbf{x},\mathbf{y},\mathbf{z})=\delta(\mathbf{y})\delta(\mathbf{z})v_c(\mathbf{x})=\delta(\mathbf{y})\delta(\mathbf{z})e^2/|\mathbf{x}| \tag{3.132}$$

in Eq. (3.25d) results in a long-range effective potential in the term $S_{\rm int}^{(1)}$, Eq. (3.47b), and only there. This long-range potential has to be screened as usual, and one should therefore add a statically screened Coulomb potential to the singlet interaction constant K_s (Finkel'stein, 1983a),

$$K_s \to K_s^{lr}=K_s-\frac{\pi}{8}(\partial n/\partial\mu)^2\frac{v_c(\mathbf{q})}{1+v_c(\mathbf{q})\chi_{\rho\rho}(\mathbf{q},\omega=0)} . \tag{3.133a}$$

Here $v_c(\mathbf{q})$ is the Fourier transform of Eq. (3.132),

$$v_c(\mathbf{q})=2^{d-1}\pi e^2/|\mathbf{q}|^{d-1},\quad d=2,3 , \tag{3.133b}$$

the factor $(1+F_0^s)^{-2}$ accounts for Fermi-liquid corrections to the vertices, and $\chi_{\rho\rho}(\mathbf{q},\omega=0)$ is the static density susceptibility (see Sec. II). In the long-wavelength limit, and for ergodic density fluctuations, we have [cf. Eq. (2.15)]

$$\chi_{\rho\rho}(|\mathbf{q}|\to 0,\omega=0)=\partial n/\partial\mu . \tag{3.133c}$$

Let us now consider the quantity $H+K_s^{lr}$. Equations (3.129a) and (3.133a) yield

$$H+K_s^{lr}=\frac{\pi}{8}\frac{\partial n}{\partial\mu}\frac{|\mathbf{q}|^{d-1}}{|\mathbf{q}|^{d-1}+\kappa_d^{d-1}} , \tag{3.133d}$$

with the screening wave number κ_d given by Eq. (2.45b). Equation (3.133d) shows that in the long-range case $H+K_s^{lr}$ is equal to the full compressibility (rather than to the screened one, which is given by $H+K_s$), which vanishes at zero wave number. Equation (3.133d) is one manifestation of the so-called compressibility sum rule (see, for example, Pines and Nozières, 1989). In the singlet propagator $\mathcal{D}^s$, Eq. (3.122f), we therefore have to keep the momentum dependence of $H+K_s^{lr}$. We conclude that one can go from a short-range interaction to the Coulomb case by making the replacements

$$K_s \to \frac{-\tilde{H}}{|\mathbf{p}|^{d-1}+\kappa_d^{d-1}}[|\mathbf{p}|^{d-1}F_0^s+\kappa_d^{d-1}(1+F_0^s)] , \tag{3.134a}$$

$$\mathcal{D}_n^s(p)\to\frac{\kappa_d^{d-1}+|\mathbf{p}|^{d-1}}{|\mathbf{p}|^{d-1}}\frac{1}{\mathbf{p}^2+|\mathbf{p}|^{3-d}\kappa_d^{d-1}+G\tilde{H}\Omega_n} , \tag{3.134b}$$

where

$$\tilde{H}=H/(1+F_0^s)=\frac{\pi}{8}\partial n/\partial\mu . \tag{3.134c}$$

Notice that the propagator shown in Eq. (3.134b) diverges for $|\mathbf{p}|\to 0$ even for nonzero ω_n. This is a direct consequence of the incompressibility of the Fermi liquid with long-range interactions. We also note that the denominator of $\mathcal{D}^s$ in the long-range case, Eq. (3.134b), has the same structure as the dynamically screened Coulomb interaction, Eq. (3.133a), with $\chi_{\rho\rho}(\mathbf{q},\omega=0)$ replaced by

$$\chi_{\rho\rho}(\mathbf{q},\omega_n)=\frac{(\partial n/\partial\mu)D\mathbf{q}^2}{\Omega_n+D\mathbf{q}^2} . \tag{3.135}$$

This shows that dynamical screening effects are correctly taken into account by Eqs. (3.134). Finally we mention that the electrical conductivity is related to the screened density susceptibility, Eq. (3.135) (see Pines and Nozières, 1989), and the Einstein relation, Eq. (2.21), still holds.

In the literature both the case of short-range interactions and that of Coulomb interactions have been studied. It has been found that both cases yield the same critical behavior for all quantities except for the single-particle DOS (Castellani, DiCastro, Lee, and Ma, 1984) and for the ultrasound attenuation (Dobrosavljevic *et al.*, 1991). In actual calculations there is a convenient criterion for the relevance or otherwise of the Coulomb interaction: it is relevant whenever the result does not allow for the limit $(H+K^s)\to 0$ to be taken.

4. Symmetry considerations

a Spontaneously broken symmetry

Let us discuss the symmetries of the field theory, first in the spinor formulation. For the sake of simplicity we do so for the free-electron action S_0, Eq. (3.43). Note, however, that the symmetry considerations given below remain valid in the presence of the random potential contribution $S_{\rm dis}$ in Eqs. (3.25) or (3.136). We split the action S_0 into two parts,

$$S_0=S_0^{\omega}+S_0^{\mathbf{k}} , \tag{3.136a}$$

where

$$S_0^{\mathbf{k}}=\sum_{\mathbf{k}}(-\mathbf{k}^2/2m+\mu)(\eta(\mathbf{k})|\eta(\mathbf{k})) , \tag{3.136b}$$

$$S_0^{\omega}=\sum_{\mathbf{k}}(\eta(\mathbf{k})|i\Omega\eta(\mathbf{k})) , \tag{3.136c}$$

with $(\eta(\mathbf{k})|\eta(\mathbf{k}))\equiv\sum_n\sum_\alpha(\eta_n^\alpha(\mathbf{k}),\eta_n^\alpha(\mathbf{k}))$, and Ω given by Eq. (3.110c). $S_0^{\mathbf{k}}$ is invariant under transformations D which leave invariant the scalar product in spinor space:

$$\begin{aligned}(D\eta|D\eta)&=(CD\eta)^TD\eta=(C\eta)^T(CDC^{-1})^TD\eta\\&=(\eta|CD^TC^{-1}D\eta)\overset{!}{=}(\eta|\eta) ,\end{aligned} \tag{3.137}$$

where we have used $C^{-1}=C^T$ with C the charge-conjugation matrix, Eq. (3.41c). The transformations D

that leave S_0^k invariant therefore obey

$$DCD^T=C, \quad C^T=-C \ . \tag{3.138}$$

Equation (3.138) identifies the D as elements of the symplectic group Sp($4N$) (see, for example, Chap. 10-18 of Hamermesh, 1962). From the definition of the composite fields $B=\eta\otimes\bar{\eta}$, Eq. (3.58), it follows that if η transforms like $\eta'=D\eta$, then B and therefore Q transform like

$$B'=D\eta\otimes\overline{D\eta}=D\eta\otimes\eta^T D^T C^T$$
$$=D\eta\otimes\eta^T C^T D^{-1}=DBD^{-1} \ , \tag{3.139a}$$

$$Q'=DQD^{-1} \ , \tag{3.139b}$$

where we have used $C^T=C^{-1}$ and Eq. (3.138). Indeed one easily checks that the first term in Eq. (3.114) is invariant[9] under $Q\rightarrow Q'$.

The second term in Eq. (3.114) breaks this symmetry, $\mathrm{tr}\Omega Q'=\mathrm{tr}\Omega' Q\neq\mathrm{tr}\Omega Q$. In the absence of interactions this symmetry breaking has been discussed in detail (McKane and Stone, 1981). In that case, the symmetry is spontaneously broken to Sp($2N$)×Sp($2N$), and the full Lagrangian has the symmetry Sp($4N$)/Sp($2N$)×Sp($2N$). The Goldstone mode associated with the spontaneous symmetry breaking is the diffusive propagator $\mathcal{D}$ ($\equiv\mathcal{D}^s\equiv\mathcal{D}^t$), whose mass is indeed proportional to the external frequency. The order parameter is given by $\langle Q_{nn}^{\alpha\beta}(\mathbf{x})\rangle_{\omega_n=i0}=N(\varepsilon_F)\equiv N_F$ with N the single-particle DOS.

In the noninteracting case, $0<N(\varepsilon)<\infty$ for all energies inside the band and for all amounts of disorder (Wegner, 1981a). The localization transition is therefore not a transition between a symmetric and a broken-symmetry phase, but rather one between a Goldstone phase and one in which localization substitutes for a Goldstone mode (McKane and Stone, 1981), with the symmetry being broken in either phase. This interpretation of the diffusive density response as a Goldstone mode resulting from the spontaneous breaking of a continuous symmetry has an interesting consequence concerning the dimensionality dependence of the metal-insulator transition. A general theorem associated with the names of Mermin and Wagner (1966) and Hohenberg (1967) in condensed-matter physics, and Coleman (1973) in field theory, states that there cannot be a Goldstone mode for $d\leq 2$. In the present context this means that there should be no diffusion and hence no metal-insulator transition in $d\leq 2$; the system will be an insulator for all nonvanishing values of the disorder. In the language of critical phenomena, $d_c^-=2$ is a lower critical dimension for the problem. This conclusion is consistent with perturbative calculations; see Sec. II.A.1.d. For $d=2+\varepsilon$, $\varepsilon\ll 1$, one expects the critical disorder to be small and related to ε. This suggests a $2+\varepsilon$ expansion for a description of the metal-insulator transition, which has been utilized in all field-theoretic approaches to the problem (see below and the following sections).

The physical significance of the interpretation of the diffusive propagator as a Goldstone mode in the noninteracting field theory is not obvious. In general there can be two distinct reasons for an excitation to be slow: an underlying conservation law or the existence of a Goldstone phase. As discussed at the beginning of Sec. II, particle number-, spin-, and energy-density fluctuations are in general slow due to the respective conservation laws.

In the presence of interactions, the third term in Eq. (3.114), $S_{\mathrm{int}}[Q]$, breaks the symmetry as well as the second term, while the DOS at the Fermi surface *does* vanish in the insulating phase; see Secs. V and VI below. Neither the mathematical nor the physical aspects of these changes in the symmetry breaking compared to the noninteracting case have been investigated. However, the presence of S_{int} does not change the fact that $d_c^-=2$, and the $2+\varepsilon$ expansion is still possible.

b. Externally broken symmetry

The Sp($4N$) symmetry can also be explicitly broken by external fields. Examples that have been discussed are external magnetic fields, magnetic impurities, and spin-orbit coupling.

Let us consider magnetic fields first. The orbital effect of an external field $\mathbf{B}(\mathbf{x})=\vec{\nabla}\times\mathbf{A}(\mathbf{x})$ is taken into account by the substitution

$$\Delta\rightarrow\left[\vec{\nabla}-\frac{ie}{c}\mathbf{A}(\mathbf{x})\right]^2 \tag{3.140a}$$

in Eq. (3.25b). By explicit calculation it is easily checked that, in the spinor formulation, Eq. (3.43), this corresponds to the substitution

$$\Delta\rightarrow\left[\vec{\nabla}-\frac{e}{c}\mathbf{A}(\mathbf{x})\tau_3\otimes s_0\right]^2 \ . \tag{3.140b}$$

Repeating the derivation of the nonlinear sigma model, Eq. (3.114), one finds that the action now reads (Efetov *et al.*, 1980)

$$S[Q]=\frac{-1}{2G}\int d\mathbf{x}\,\mathrm{tr}\left[\vec{\nabla}\tilde{Q}(\mathbf{x})+\frac{e}{c}\mathbf{A}(\mathbf{x})[\tilde{\mathbf{Q}}(\mathbf{x}),\tau_3\otimes s_0]\right]^2$$
$$+2H\int d\mathbf{x}\,\mathrm{tr}(\Omega\tilde{Q}(\mathbf{x}))+S_{\mathrm{int}}[\tilde{Q}] \ . \tag{3.141}$$

Apart from the orbital term, there is a Zeeman term in the Hamiltonian, which in first quantization reads $-g_L\mu_B\frac{1}{2}\vec{\sigma}\cdot\mathbf{B}$. Here $g_L\simeq 2$ is the Landé factor, $\mu_B=1/2mc$ is the Bohr magneton, and $\vec{\sigma}=(\sigma_1,\sigma_2,\sigma_3)$ with σ_i the Pauli matrices. In Eq. (3.25a) this contributes a term

[9]Notice that the symmetry group here is the compact group Sp($4N$), while in Wegner's original model (Wegner, 1979) it was the noncompact group $O(2N,2N)$. Höf and Wegner (1986) have shown that nonlinear sigma models with compact symplectic and noncompact pseudo-orthogonal symmetries, respectively, have identical perturbation expansions.

$$S_B = -\tfrac{1}{2} g_L \mu_B B \int dx \sum_\sigma \bar{\psi}_\sigma(x)(-)^\sigma \psi_\sigma(x) \ . \tag{3.142a}$$

In bispinor language this corresponds to

$$S_B = -\tfrac{1}{2} g_L \mu_B B \int dx \sum_\alpha (\eta^\alpha(x), \tau_3 \otimes s_3 \eta^\alpha(x)) \ , \tag{3.142b}$$

and in Eq. (3.114) this adds a term (Finkel'stein, 1984a, 1984b)

$$S_B[Q] = b \int d\mathbf{x} \, \mathrm{tr}(\tau_3 \otimes s_3 Q(\mathbf{x})) \ , \tag{3.143}$$

where $b = \pi N_F g_L \mu_B B / 2$.

Now we turn to magnetic impurities. In the Hamiltonian there is a term $\mathbf{u}_s(\mathbf{x}) \cdot \vec{\sigma}$, where $\mathbf{u}_s(\mathbf{x})$ is a random potential. As for the spin-independent random scattering potential, we assume a Gaussian distribution [cf. Eq. (3.27)], with second moment

$$\{u_s^i(\mathbf{x}) u_s^j(\mathbf{y})\} = \frac{1}{6\pi N_F \tau_s} \delta_{ij} \delta(\mathbf{x} - \mathbf{y}) \ , \tag{3.144}$$

where τ_s is the magnetic scattering mean free time. In Eq. (3.25a) this contributes a term

$$S_s = i \int dx \, \bar{\psi}^T(x) \sum_{j=1}^{3} u_s^j(\mathbf{x}) s_j \psi(x) \ . \tag{3.145a}$$

In bispinor language this corresponds to

$$S_s = i \int dx \sum_{j=1}^{3} (-)^{j+1} u_s^j(\mathbf{x}) \sum_\alpha (\eta^\alpha(x), \tau_3 \otimes s_j \eta^\alpha(x)) \ . \tag{3.145b}$$

In Eq. (3.114) this adds a term (Efetov *et al.*, 1980)

$$S_s[Q] = -\frac{\pi N_F}{6\tau_s} \int d\mathbf{x} \sum_{i=1}^{3} \mathrm{tr}(\Sigma_i Q(\mathbf{x}))^2 \ , \tag{3.146a}$$

where

$$(\Sigma_i)^{\alpha\beta}_{nm} = \delta_{\alpha\beta} \delta_{nm} \tau_3 \otimes s_i \ . \tag{3.146b}$$

Finally, we consider the spin-orbit interaction. In the Hamiltonian, there is a term (Davydov, 1965, p. 252) $-ic\vec{\sigma} \cdot [\vec{\nabla}(v_{so}(\mathbf{x}) + u_{so}(\mathbf{x})) \times \vec{\nabla}]$ with a constant c. Here $v_{so}(\mathbf{x})$ is the spin-orbit interaction in the absence of impurities, and u_{so} is the impurity spin-orbit interaction. This leads to an additional term in Eq. (3.114) (Efetov *et al.*, 1980),

$$S_{so}[Q] = \frac{\pi N_F}{2} \int d\mathbf{x} \sum_{i=1}^{3} \frac{1}{\tau_{so}^i} \mathrm{tr}(s_i Q(\mathbf{x}))^2 \ , \tag{3.147}$$

where τ_{so}^i is the spin-orbit scattering time in spin channel number i. All of these additional interactions are assumed to be sufficiently weak that they can be considered independent of each other and of elastic impurity scattering. In this case, the respective contributions to the action simply add.

c. Remaining soft modes and universality classes

The various additional terms in the action which appear in the presence of magnetic fields, magnetic scattering, or spin-orbit scattering lead to some of the propagators given in Eqs. (3.123) acquiring a mass, i.e., they are finite even at zero momentum and frequency. With the help of Eq. (3.117a) or (3.117c), and keeping only terms quadratic in q or w, respectively, one finds the following (Efetov *et al.*, 1980). The orbital effect of the magnetic field, Eq. (3.141), produces a mass proportional to B in the particle-particle channels, $r = 1, 2$, irrespective of the spin.[10] The Zeeman term, Eq. (3.142a), furthermore produces a mass in two ($i = 1, 2$) of the three spin-triplet channels (Finkel'stein, 1984a). Magnetic impurities result in a mass proportional to $1/\tau_s$ in the particle-particle channel, and additionally in all particle-hole spin-triplet channels. Finally, spin-orbit scattering results in a mass proportional to $1/\tau_{so}$ in all spin-triplet channels. These mass terms change the effective model, since according to the soft-mode philosophy explained in Sec. II and earlier in the present section, the massive modes can be discarded if one is interested in asymptotic low-temperature properties.[11] We thus have to distinguish between the following cases, which give rise to different universality classes for the metal-insulator transition.

(1) *Weak magnetic field.* In this case, which is the one originally studied by Finkel'stein (1983a), one assumes a magnetic field strong enough to suppress the particle-particle channel, but weak enough so that the Zeeman term can be neglected. In a real system, this case can at most be realized in a transient temperature region. The particle-hole propagators are given by Eq. (3.122b), while the particle-particle propagators, Eq. (3.122c), are suppressed. Since ${}_{1,2}q$ are irrelevant, the matrix elements of ${}^i q$ (or w) are complex numbers rather than quaternions. The expansion (3.117b) can still be used if r is restricted to $r = 0, 3$ with $\tau_0 = 1$, $\tau_3 = i$.

(2) *Strong magnetic field.* In this case the Zeeman splitting is taken into account (Castellani, DiCastro, Lee, and Ma, 1984; Finkel'stein, 1984a). Here, in addition to the restrictions of case (1), the spin channels are restricted to $i = 0, 3$ (i.e., the z component of the spin must van-

[10] This is not the only effect of a magnetic field. In the action, Eq. (3.141), a term has been left out which contains three Q matrices and two gradients, and whose coupling constant is given by the Hall conductivity (Pruisken, 1984). In fields strong enough to lead to Landau quantization of the energy levels, and in the absence of interactions, this term gives rise to the integer quantum Hall effect (von Klitzing *et al.*, 1980; for a review, see Prange and Girvin, 1987). It has also been used to analyze the scaling behavior of the Hall conductivity; see Sec. III.C below. With interactions, the corresponding effect is generally believed to be the fractional quantum Hall effect (see, for example, Chakraborty and Pietiläinen, 1988). This has not been studied in the framework of the nonlinear sigma model, and we shall ignore the case of a quantizing magnetic field.

[11] For small masses, one should actually keep them and study the crossover phenomena associated with these terms. For the case of interacting electrons, this has not been done so far.

TABLE I. Universality classes for the metal-insulator transition.

Symbol	Symmetry breaker	Diffusive modes	Interaction
MF(SR)	magnetic field	$r=0,3$ $i=0,3$	short-range
MF(LR)			Coulomb
MI(SR)	magnetic impurities	$r=0,3$ $i=0$	short-range
MI(LR)			Coulomb
SO(SR)	spin-orbit scattering	$r=0,1,2,3$ $i=0$	short-range
SO(LR)			Coulomb
G(SR)	none	all	short-range
G(LR)			Coulomb

ish). The propagators for r, $i=0,3$, are still given by Eq. (3.122b), and the ${}^{i}q$ are complex numbers.

(3) *Magnetic impurities.* Here all of the diffusive modes are suppressed, except for the particle-hole spin-singlet channel. The latter propagator ($r=0,3$, $i=0$) is still given by Eq. (3.122b), and the ${}^{i=0}q_{nm}^{\alpha\beta}$ are complex numbers.

(4) *Spin-orbit coupling.* Here all spin-singlet modes ($i=0$) and only these are diffusive. The corresponding propagators are given by Eqs. (3.122).

(5) *Generic case.* This is the case in which no external symmetry-breaking fields are present. All diffusive modes are present, and the propagators are given by Eqs. (3.122).

We shall see in Sec. V that all of these cases result in different critical behavior, depending on whether the electrons interact via the long-range Coulomb interaction or via a short-range interaction. Excluding case (1), which is realizable only in a transient regime, we then have eight universality classes for metal-insulator transitions that can be realized, at least in principle, in actual systems.[12] They are summarized in Table I. In Secs. IV and VI we shall see that the generic case also allows for a magnetic phase transition which is distinct from the metal-insulator transition.

C. Summary of results for noninteracting electrons

Before we describe results obtained for the full model derived above, let us briefly summarize what is known about the nonlinear sigma model for noninteracting electrons, i.e., Eqs. (3.114)–(3.116) with $S_{\rm int}=0$. In the terminology of the last section, the generic case, the spin-orbit coupling case, and the cases of weak magnetic fields or magnetic impurities are described by models with $\mathrm{Sp}(n)/\mathrm{Sp}(n-p)\times\mathrm{Sp}(p)$, $\mathrm{U}(n)/\mathrm{U}(n-p)\times\mathrm{U}(p)$, and $\mathrm{O}(n)/\mathrm{O}(n-p)\times\mathrm{O}(p)$ symmetries, respectively (Wegner, 1979; Efetov *et al.*, 1980; Hikami, 1980; Hikami *et al.*, 1980; Schäfer and Wegner, 1980), in the limit[13] $n=p=0$.

1. Renormalization of the nonlinear sigma model

All of the above-mentioned nonlinear sigma models are renormalizable with two renormalization constants. This was shown by Brézin *et al.* (1976) for the $\mathrm{O}(n)$ vector model, and their proof applies to the matrix models as well. The two renormalization constants needed are a field renormalization constant Z and a coupling constant renormalization constant Z_g. In the replica limit, one finds $Z=1$, order by order, in the loop expansion. The physical meaning of this result is that the density of states is not critical at the metal-insulator transition (Wegner, 1981a). Z_g connects the bare resistivity G with its renormalized counterpart g via

$$G=\kappa^{-\varepsilon}Z_g g\ , \tag{3.148}$$

where $\varepsilon=d-2$, and κ is the arbitrary momentum scale of the RG. Z_g determines the β function

$$\beta(g)=\frac{dg}{d\ln\kappa}=\frac{\varepsilon g}{1+g(d\ln Z_g/dg)}\ . \tag{3.149}$$

The β function has been determined by a direct calcula-

[12]For most experimental systems the Coulombic case is relevant. However, Castellani, DiCastro, Lee, and Ma (1984) and Altshuler, Aronov, and Zuzin (1982) have argued that for ^{3}He on a disordered substrate and for electrons in certain semiconductors devices the effective interaction is short ranged.

[13]This is the description in terms of compact groups, which arises from a formulation in terms of Grassmannian fields (Efetov *et al.*, 1980). In the noninteracting case, one can also start from bosonic fields (Wegner, 1979). One then finds a description in terms of noncompact groups, $\mathrm{O}(n-p,p)$, etc., and the physical roles of orthogonal and symplectic symmetries are interchanged.

tion to four-loop order (Wegner, 1989). Hikami (1990) recently noticed that the nonlinear sigma models in question are closely related to certain classes of string theories. Using this relation, he confirmed Wegner's four-loop results and obtained the five-loop-order contributions (Hikami, 1992). To this order the results for the symplectic (s), unitary (u), and orthogonal (o) cases are

$$\beta_s(g)=\varepsilon g-g^2-\tfrac{3}{4}\zeta(3)g^5+\tfrac{27}{64}\zeta(4)g^6+O(g^7)\ , \tag{3.150a}$$

$$\beta_u(g)=\varepsilon g-2g^3-6g^5+O(g^7)\ , \tag{3.150b}$$

$$\beta_0(g)=\varepsilon g+2g^2-12\zeta(3)g^5-\tfrac{27}{2}\zeta(4)g^6+O(g^7)\ . \tag{3.150c}$$

All three β functions allow for a fixed point g^*, $\beta(g^*)=0$, which corresponds to a metal-insulator transition. The spin-orbit coupling case is special in that Eq. (3.150c) seems to suggest a metal-insulator transition with nonzero g^* in $d=2$. The β functions for the symplectic and the orthogonal cases are related, and in the unitary case only odd powers of g appear (Wegner, 1981b). Equations (3.150) yield the localization length exponent $\nu=-1/\beta'(g^*)$ and the conductivity exponent $s=\nu\varepsilon$ to five-loop order. Wegner (1986, 1989) has also determined the crossover exponents from symplectic to unitary or orthogonal behavior. While the four-loop terms in Eqs. (3.149) yield large negative corrections to the one-loop result $\nu=1/\varepsilon$, a Borel-Padé approximation based on the five-loop result for the symplectic case yields a value for ν that respects the rigorous inequality $\nu\geq 2/d$ (Chayes *et al.*, 1986). Still, the unknown convergence properties of the ε expansion have given rise to the suggestion that the nonlinear sigma model may not be complete. It has been found that higher-order gradient terms acquire large positive corrections to their bare dimensions and thus may be relevant, contrary to what one assumes in deriving the nonlinear sigma model (Altshuler *et al.*, 1988; Kravtsov *et al.*, 1988; Lerner and Wegner, 1990; Wegner, 1990). A possible consequence is a breakdown of the ε expansion. These investigations are not quite conclusive yet, and further work on this subject is needed. We shall come back to this subject in Sec. X.

From the loop expansion for the O(n) vector model, an equation of state can be constructed which gives results in the high-temperature phase as well as in the low-temperature phase and in the critical region (Brézin and Zinn-Justin, 1976; Nelson and Rudnick, 1976). In the localization problem, the "equation of state" takes the form of a transcendental equation for the dynamical conductivity, which yields a physical solution in the insulating phase (Hikami, 1982; Pruisken and Wang, 1989). The "equation of state" technique amounts to a resummation of the loop expansion, which allows one to go beyond perturbation theory. This can be done in a systematic way, yielding successive approximations for the "equation of state" based on higher and higher orders in the loop expansion (Belitz and Yang, 1993). For the symplectic model, the one-loop "equation of state" coincides with results obtained earlier by means of mode-coupling techniques (Belitz *et al.*, 1981; Vollhardt and Wölfle, 1982; Wölfle and Vollhardt, 1982).

2. Composite operators, ultrasonic attenuation, and the Hall conductivity

Among all correlation functions, the conductivity plays a special role, since it appears as a coupling constant in the nonlinear sigma model. In order to study other correlation functions, e.g., the stress correlation that determines the ultrasonic attenuation (see Sec. II.A.2.b), one has to add an appropriate operator to the model. This was first studied by Wegner (1980), who calculated the critical behavior of the participation ratio this way. Later he gave a complete classification of all composite scaling operators with no spatial derivatives and calculated their anomalous dimensions (Wegner, 1987a, 1987b). The critical behavior of the transverse ultrasonic attenuation α was found to be given as the sum of two such scaling operators, one of which is relevant while the other is irrelevant (Castellani and Kotliar, 1986, 1987; Kirkpatrick and Belitz, 1986b):

$$\alpha(\omega,t)/\alpha^0(\omega)=b^{x_+}\alpha_+(\omega b^z,tb^{1/\nu})+b^{x_-}\alpha_-(\omega b^z,tb^{1/\nu})\ . \tag{3.151a}$$

Here $\alpha^0(\omega)\propto\omega^2$ is the Boltzmann result; see Eqs. (2.51). b is an arbitrary scale factor, $t=|g^*-g|/g^*$ is the dimensionless distance from the critical point, ν is the localization length exponent, and $z=d$ for noninteracting electrons. The anomalous dimensions x_+ and x_- in the symplectic case are

$$x_+=2\varepsilon-\tfrac{3}{2}\zeta(3)\varepsilon^4+O(\varepsilon^5)\ , \tag{3.151b}$$

$$x_-=-\varepsilon+\tfrac{3}{2}\zeta(3)\varepsilon^4+O(\varepsilon^5)\ , \tag{3.151c}$$

and in the unitary case are

$$x_\pm=\pm\sqrt{2\varepsilon}+O(\varepsilon)\ . \tag{3.152}$$

On approaching the critical point at zero frequency or wave number, α/α^0 diverges like

$$\alpha/\alpha^0\sim t^{-\nu x_+}\ . \tag{3.153a}$$

For the critical dynamics at $t=0$ one finds

$$\alpha(\omega)/\alpha^0(\omega)\sim\omega^{-x_+/d}\ . \tag{3.153b}$$

In an expansion in powers of $(\varepsilon\ln\omega)$, both the relevant and the irrelevant operator contribute. This is the reason underlying the observation (Kirkpatrick and Belitz, 1986a) that the critical behavior cannot simply be obtained by an exponentiation of perturbation theory.

A similar phenomenon has been shown to occur in the case of the Hall conductivity (Wang *et al.*, 1992). Again, perturbation theory has a contribution from an irrelevant operator, and simple exponentiation does not give the correct critical behavior. The action for noninteracting

electrons in a magnetic field is given by Eq. (3.141) with $S_{\rm int}=0$ and $G=4/\pi\sigma^0_{xx}$, where σ^0_{xx} is the diagonal element of the conductivity tensor in the self-consistent Born approximation. In addition, there is a contribution (Pruisken, 1984; Biafore *et al.*, 1990)

$$S_{xy}[Q]=\frac{-i}{2G_{xy}}\int d\mathbf{x}\sum_{\mu,\nu}\varepsilon_{\mu,\nu}\mathrm{tr}[Q\nabla_\mu Q\nabla_\nu Q(\tau_3\otimes s_0)]\ , \tag{3.154}$$

with $G_{xy}=4/\pi\sigma^0_{xy}$, and $\varepsilon_{\mu\nu}$ the antisymmetric tensor of rank two. Finally, Wang *et al.* (1992) have shown that for a proper scaling description of the Hall conductivity σ_{xy} one also has to keep a term with four gradients,

$$S_\lambda[Q]=-i\lambda\int d\mathbf{x}\sum_{\mu,\nu}\mathrm{tr}(\nabla_\mu Q\nabla_\mu Q\nabla_\nu Q\nabla_\nu QQ)\ , \tag{3.155}$$

where $\lambda\sim k_F^{d-2}l^2$ with l the mean free path. Even though λ has an engineering dimension of d-4 and therefore is irrelevant near two dimensions, it was shown to couple into the flow equation for σ_{xy} in the form of the combination λB, whose scaling dimension is d-2, and therefore it has to be kept. As a result (Wang *et al.*, 1992), both σ_{xx} and σ_{xy} vanish at the transition as

$$\sigma_{xx}(t)\sim\sigma_{xy}(t)\sim t^{\varepsilon\nu}\ , \tag{3.156a}$$

and the Hall coefficient diverges,

$$R_H=\sigma_{xy}/\sigma^2_{xx}B\sim t^{-\varepsilon\nu}\ . \tag{3.156b}$$

3. Conductance fluctuations

The technique used to calculate the ultrasonic attenuation has also been applied for a direct calculation of the conductivity. Castellani and Kotliar (1987) have found that the conductivity, in contrast to the ultrasonic attenuation, is given in terms of a local operator (with no gradients) and a nonlocal one which contains a gradient of the Q matrix. The explicit calculation (Castellani and Kotliar, 1987; Pruisken and Wang, 1989) recovers the result obtained from renormalizing the nonlinear sigma model.

This direct calculation of the conductivity, or the conductance, has been extended to study conductance fluctuations within the framework of the nonlinear sigma model (Altshuler *et al.*, 1986; DeSouza and Kirkpatrick, 1991). This work confirmed earlier results obtained from many-body perturbation theory (Lee *et al.*, 1987) and established that the distribution of fluctuations of the conductance and the density of states is not purely Gaussian, but has log-normal tails that get stronger with increasing disorder. The latter conclusion was reached in the framework of an extended sigma model, with the higher-order gradient operators referred to above taken into account. The present status of this work has been summarized recently by Altshuler *et al.* (1991).

IV. SCALING SCENARIOS FOR THE DISORDERED ELECTRON PROBLEM

In the following two sections we shall show how the perturbative RG can be used to obtain explicit solutions of the field theories derived and discussed in Sec. III. This explicit solution procedure is quite complicated. Before addressing it, we discuss some general properties of the metal-insulator transition which one expects assuming that it is a conventional zero-temperature continuous phase transition.[14] In particular, we examine some of the phase-transition scenarios that general scaling theory suggests. In subsequent sections we shall explicitly verify that each of the scenarios that can be constructed on general grounds is realized in the explicit RG solution of the field theories. From this same general viewpoint we shall then discuss the possibility that disorder can cause phase transitions that are separate from the metal-insulator transition. In Secs. VI and VIII we shall see that this situation can also occur. The scaling picture that will be developed in this section is for the most part a summary and coherent exposition of results obtained by Wegner (1976), Lee (1982), Castellani, DiCastro, Lee, and Ma (1984), Finkel'stein (1984a), Abrahams and Lee (1986), Castellani and DiCastro (1986), Castellani, Kotliar, and Lee (1987), Kirkpatrick and Belitz (1990a, 1992b), Raimondi *et al.* (1990), and Belitz and Kirkpatrick (1991).

Before beginning our specific discussion of the metal-insulator transition we point out some general features of phase transitions (see, for example, Ma, 1976; Fisher, 1983). First we consider the equilibrium properties of conventional phase transitions at finite temperatures. These are determined by two independent critical exponents. All other critical exponents are related to the two independent ones by scaling laws or equalities. These two exponents characterize how the system moves away from the critical surface under scaling. If T_c is the critical temperature, then the distance from the critical surface is given by the reduced temperature $t=(T-T_c)/T_c$ and by the external field h, which is conjugate to the order parameter describing the phase transition. Let $\mu=(t,h)$ define the parameter space. Under RG iterations, the system moves away from criticality according to

$$\mu\to\mu(b)=(b^{1/\nu}t,b^{y_h}h)\ . \tag{4.1a}$$

Here ν is the correlation length exponent and b is the RG length rescaling factor. The conjugate field exponent

$$y_h=(d+2-\eta)/2 \tag{4.1b}$$

is related to the exponent η which gives the critical wave-number dependence of the order-parameter suscep-

[14] By "conventional" we mean, for instance, that there is a stable RG fixed point with one relevant operator at zero field and temperature and one relevant diverging length scale.

tibility [see Eq. (4.9c) and Table II below]. Other physical parameters distinguishing different systems are irrelevant in the RG sense. The dependence of μ on b in turn implies that the singular part of any thermodynamic quantity, $Q(t,h)$, near the critical point is a generalized homogeneous function of second degree and scales as

$$Q(t,h)=b^{x_Q}Q(b^{1/\nu}t,b^{y_h}h)\ , \tag{4.2a}$$

with x_Q depending on Q. Similarly, an equilibrium correlation function G, at wave vector $\mathbf{k}$, scales as

$$G(\mathbf{k};t,h)=b^{x_G}G(b\mathbf{k};b^{1/\nu}t,b^{y_h}h)\ . \tag{4.2b}$$

For a complete discussion of finite-temperature equilibrium phase transitions one also needs to consider equilibrium time correlation functions (see, for example, Hohenberg and Halperin, 1977). Near a critical point one expects divergent relaxation times and, in addition to ν and y_h, a dynamical scaling exponent z is needed. If we characterize the divergent time scale by $\tau\sim\xi^z$, with ξ the correlation length, then a frequency-dependent time correlation function $C(\mathbf{k},\Omega;t,h)$ scales as

$$C(\mathbf{k},\Omega;t,h)=b^{x_C}C(b\mathbf{k},b^z\Omega;b^{1/\nu}t,b^{y_h}h)\ . \tag{4.2c}$$

For zero-temperature quantum phase transitions such as the metal-insulator transition the situation is slightly more complicated (Hertz, 1976). In this case static and dynamic quantities cannot be separated, and in general any quantity will depend on the three exponents ν, y_h, and z. Similarly, the scaling equalities will depend on these three independent exponents. Physically, z is needed to describe quantum phase transitions because it reflects the role of temperature as a relevant operator that has the same scaling behavior as frequency.

Finally, a general point of experimental relevance should be emphasized. Scaling is valid only in a small region of parameter space around the critical point known as the critical region, and there are systematic corrections to scaling (Wegner, 1972). For standard thermal phase transitions, the size of the critical region is at most $t\lesssim 10^{-2}$, and often it is much smaller. For an accurate experimental determination of the critical exponents, the following conditions are necessary (see, for example, Levelt Sengers and Sengers, 1981): (i) the measurements must cover several decades of t, and (ii) the experiment must be in the critical region, or (ii′) corrections to scaling must be taken into account in analyzing the data. Because of these stringent conditions, precision measurements of critical exponents are difficult. For thermal phase transitions they are nevertheless possible, since t is relatively easy to control. The situation is less favorable at the zero-temperature metal-insulator transition in which t is related to the disorder, or possibly to a combination of disorder and interaction strength. Even in the most accurate experiments to date (Rosenbaum *et al.*, 1983) values of t smaller than 10^{-3} have not been achieved. Furthermore, the question of corrections to scaling at the metal-insulator transition has been ignored by both theory and experiment. Consequently the size of the asymptotic critical region is not known, and it is not clear whether the measured "exponents" really represent asymptotic critical exponents or rather some preasymptotic effective exponents. This is an important point, which will be discussed in detail in Secs. V and VI.

A. Metal-insulator phase-transition scenarios

1. The noninteracting problem

Consider the noninteracting field theory given by Eq. (3.118) with $S_{\rm int}[\tilde{Q}]=0$. The two parameters that occur in the field theory are G and H. Because the frequency appears in the second term in Eq. (3.118), the dynamical scaling exponent z is related to how H renormalizes. The disorder G is analogous to the temperature in finite-temperature phase transitions, and the behavior of G under scaling can be related to the correlation length exponent ν. The order parameter for the metal-insulator transition is the single-particle DOS at the Fermi level, N_F, which is given by the expectation value of ${}^0_0Q^{\alpha\alpha}_{nn}$ [see Eq. (3.131)]. For the noninteracting problem only elastic electron-impurity collisions are taken into account, and consequently there is no frequency mixing. This implies that H is the field conjugate to the order parameter. The renormalization of H therefore gives y_h as well as z.

Using the fact that ξ is the relevant length in the problem, we note for future use that the scale dimensions of G and H are

$$[G]=2-d=-\varepsilon\ , \tag{4.3a}$$

$$[H]=d-z\ . \tag{4.3b}$$

Here $\varepsilon=d-2$, and we follow Ma (1976) in defining the scale dimension of ξ as -1. The exponent z can be determined by using the fact that N_F is not sensitive to the metal-insulator transition, i.e., that the order parameter is not critical. This has been proven rigorously by Wegner (1981a), and McKane and Stone (1981) have given a detailed discussion of this aspect of the metal-insulator transition; see also the discussion in Sec. III.B.4.a. In the framework of our discussion it is a consequence of N_F's being related to the single-particle Green's function, which does not depend on diffusion corrections in the noninteracting limit. In any case, this implies that H in Eq. (3.118) is not renormalized. Its scale dimension therefore vanishes, and Eq. (4.3b) gives

$$z=d=y_h\ . \tag{4.4}$$

If we denote the dimensionless renormalized disorder by g, the dimensionless distance from the critical surface at zero temperature is $t=(g^*-g)/g^*$, with g^* the fixed-point value of g. As in Eq. (4.1a), $t(b)$ grows under RG iterations with the correlation length exponent ν,

$$t(b)=b^{1/\nu}t\ . \tag{4.5}$$

The physical quantity of interest near the metal-insulator transition is the electrical conductivity $\sigma(t,T)$. In what follows, temperature T and frequency Ω will be used interchangeably. Presumably this is correct in a scaling sense. In our units the conductivity has the dimension of a length to the power $(2-d)$. Since we assume that ξ is the only relevant length, a critical σ naively implies $\sigma \sim 1/\xi^{d-2}$. This further implies that the frequency- and wave-number-dependent electrical conductivity satisfies the scaling law

$$\sigma(t,\mathbf{k},\Omega)=b^{-\varepsilon}\sigma(b^{1/\nu}t,b\mathbf{k},b^{z}\Omega) . \tag{4.6a}$$

See Eq. (1.3). For the noninteracting problem, the naive argument leading to the exponent $-\varepsilon$ in Eq. (4.6a) can be made rigorous (Wegner, 1976, 1981a). For the interacting problem this is not the case, and to obtain the most general scaling scenario one should replace ε in Eq. (4.6a) by $\varepsilon+\vartheta$ with ϑ an unknown critical exponent. This implies that, to obtain the general scaling theory for metal-insulator transition, one should replace ε in the equations below by $\varepsilon+\vartheta$. It should be stressed, however, that in the context of the nonlinear-sigma-model description of disordered electronic systems, $\vartheta=0$ unless either the disorder (Castellani, Kotliar, and Lee, 1987) or the electric charge (Belitz and Kirkpatrick, 1993) is a dangerous irrelevant variable. So far there have been no consistent theoretical calculations that give $\vartheta\neq 0$ for systems undergoing a metal-insulator transition, and in what follows we consider only the case in which $\vartheta=0$. We shall come back to this point in connection with Eq. (4.16b) below. The Einstein relation between σ and the charge diffusion coefficient D_c [see Eqs. (2.21) or (3.128a), (3.129a)] yields

$$D_c(t,\mathbf{k},\Omega)=b^{-\varepsilon}D_c(b^{1/\nu}t,b\mathbf{k},b^{z}\Omega) . \tag{4.6b}$$

Note that Eq. (4.6a) implies that at $\mathbf{k}=0$, $\Omega=0$,

$$\sigma(t\to 0)\sim t^{s} , \tag{4.6c}$$

with,

$$s=\nu(d-2) . \tag{4.6d}$$

Equation (4.6d) relates the correlation length exponent ν to the conductivity exponent s. It is known as Wegner's scaling law. In the absence of dangerous irrelevant variables it is expected to hold at all metal-insulator transitions (a definition and discussion of dangerous irrelevant variables can be found in the lectures by Fisher, 1983. See also Ma, 1976, Sec. VII.4).

We next derive several scaling equalities for zero-temperature phase transitions and show that Eq. (4.4) and the arguments leading to it are consistent with these identities. We start with the free energy per unit volume,

$$f(t,T,h)=-\frac{1}{\beta V}\log Z , \tag{4.7}$$

with $\beta=T^{-1}$ the transverse temperature, V the system volume, Z the partition function, and h the field conjugate to the order parameter. According to our previous discussion, f satisfies the scaling equation

$$f(t,T,h)=b^{-(d+z)}f(b^{1/\nu}t,b^{z}T,b^{y_h}h) . \tag{4.8}$$

The order-parameter density n_0 satisfies

$$\begin{aligned} n_0(t,T)&=-\frac{1}{T}\frac{\partial}{\partial h}\bigg|_{h=0} f(t,T,h) \\ &=b^{-(d-y_h)}n_0(b^{1/\nu}t,b^{z}T) . \end{aligned} \tag{4.9a}$$

For the critical exponent β this implies

$$\beta=\nu(d-y_h)=\frac{\nu}{2}(d+\eta-2) , \tag{4.9b}$$

where we have used Eq. (4.1b). Notice that in giving the first equality in Eq. (4.9a) we introduced a factor T^{-1}, which cancels the T in Eq. (4.7). In addition we have assumed that h and not h/T is the field conjugate to the order parameter. The physical motivation for these definitions is that the order parameter for the metal-insulator transition is local in frequency space and not in imaginary-time space. The situation is in general different at other quantum phase transitions (cf. Fisher *et al.*, 1989), and consequently the scaling descriptions of such transitions can appear different, although they are physically equivalent. Similarly the order-parameter fluctuations define the order-parameter susceptibility χ_0, which is equal to the order-parameter density correlation function G_0 at zero wave number,

$$\chi_0(t,T)=\frac{1}{T}\frac{\partial^2}{\partial h^2}\bigg|_{h=0} f(t,T,h)=G_0(\mathbf{k}=0;t,T) . \tag{4.9c}$$

Here χ_0 obeys the scaling law

$$\chi_0(t,T)=b^{(2y_h-d)}\chi_0(b^{1/\nu}t,b^{z}T) , \tag{4.9d}$$

which implies for the critical exponent γ

$$\gamma=\nu(2y_h-d)=\nu(2-\eta) . \tag{4.9e}$$

For later reference we note that Eq. (4.9c) with $\mathbf{k}\neq 0$ defines the wave-vector-dependent order-parameter susceptibility.

From Eq. (4.9b) we see that $y_h=d$ is consistent with a noncritical order parameter at the metal-insulator transition. $\beta=0$ also gives $\eta=2-d$. To establish the consistency of this, we note that the only propagator that appears in the noninteracting limit is the density-density correlation function, Eq. (2.36); see also Eq. (3.122e). It reads

$$\Phi^{R}_{\rho\rho}(\mathbf{k},\Omega)=\frac{i\partial n/\partial\mu}{-i\Omega+D_c(t,\mathbf{k},\Omega)\mathbf{k}^2} . \tag{4.10}$$

With the help of Eq. (4.6b), we see that this is indeed consistent with $\eta=2-d$ (Abrahams and Lee, 1986). Note that Eqs. (4.9d) and (4.9e) and $y_h=d$ imply that the order-parameter susceptibility is divergent at the noninteracting metal-insulator transition. This same argument

can be used for the interacting metal-insulator transition, which will be discussed in the next subsection. For a consistent scaling description in that case one needs to take into account the fact that the physical propagator is a renormalized propagator times a wave-function renormalization factor squared. The wave-function renormalization can be related to the critical behavior of the order parameter. The absence of a wave-function renormalization in the noninteracting case is equivalent to the order parameter's being uncritical, as was already mentioned in Sec. III.C.1. We stress again that the noninteracting metal-insulator transition is unusual in a scaling sense only because the order parameter is not critical.

Finally, we note for future use that Eq. (4.8) implies a scaling law for the specific heat (Castellani and DeCastro, 1986),

$$C_v(t,T)=-T\frac{\partial^2}{\partial T^2}f(t,T,0)=b^{-d}C_v(b^{1/\nu}t,b^z T) \\ =b^{z-d}T\gamma(b^{1/\nu}t,b^z T)\ . \tag{4.11}$$

Notice that, for noninteracting systems, $z=d$ and Eq. (4.11) gives $C_v(T\to 0)\propto T$, i.e., standard analytic Fermi-liquid behavior, even at the critical point. In giving the final equality in Eq. (4.11) we have defined a function $\gamma(t,T)$ which is a generalization of the usual specific-heat coefficient. We have also used the fact that $C_v(T\to 0)$ must vanish according to the third law of thermodynamics.

2. The interacting problem

The interacting field theories are given by Eqs. (3.118) and (3.92). The parameters that occur in the field theory are G, H, and $K^{(s)}$, and possibly $K^{(t)}$ and $K^{c(s)}$, depending on the universality class, as discussed in Sec. III.B.4.c. As compared to the noninteracting problem, there are some conceptual problems. Examining Eqs. (3.122e), (3.122f), and (3.122g), and scaling the wave number with ξ, gives three time scales, which in principle all diverge differently at the metal-insulator transition (Kirkpatrick and Belitz, 1990a). If g is the dimensionless renormalized disorder and h, k_s, and k_t are the renormalized values of H, $K^{(s)}$, and $K^{(t)}$, then the three time scales and dynamical scaling exponents are

$$\tau_1=\xi^d g h = A_1\xi^{z_1}\ , \tag{4.12a}$$

$$\tau_2=\xi^d g h(1+\gamma_t)= A_2\xi^{z_2}\ , \tag{4.12b}$$

$$\tau_3=\xi^d g h(1+\gamma_s)= A_3\xi^{z_3}\ , \tag{4.12c}$$

with $\gamma_{s,t}=k_{s,t}/h$. These are in general three different time scales. They coincide if and only if γ_s and γ_t are finite constants not equal to -1 at the phase transition. The possibility of multiple divergent time scales at a critical point has been discussed before for finite-temperature phase transitions. Fixed points with multiple divergent relaxation times have been dubbed weak dynamical scaling fixed points (De Dominicis and Peliti, 1978). For future use we also note that the zero-temperature RG flow equations for g, γ_s, γ_t, and $\gamma_c=k_c/h$ (k_c is the renormalized value of $K^{c(s)}$) can depend only on g, γ_s, γ_t, and γ_c, i.e., only ratios of time scales will appear.

A simple way to interpret the meaning of multiple divergent time scales is this: if $Q(\Omega)$ is a scaling function of Ω, then it is natural to assume that

$$Q(\Omega)=b^{x_Q}Q(b^{z_1}\Omega,b^{z_2}\Omega,b^{z_3}\Omega)\ . \tag{4.13a}$$

Suppose $z_1>z_2>z_3$ and let $b=\Omega^{-1/z_1}$. Then

$$Q(\Omega)=\Omega^{-x_Q/z_1}Q(1,\Omega^{1-z_2/z_1},\Omega^{1-z_3/z_1})\ . \tag{4.13b}$$

Near the critical point, $\Omega\to 0$, one has

$$Q(\Omega)=\Omega^{-x_Q/z_1}Q(1,0,0)\ . \tag{4.13c}$$

From Eq. (4.13c) one might conclude that the subdominant divergent time scales are irrelevant and that *the* dynamical scaling exponent is z_1. We shall see, however, that sometimes this argument breaks down because $Q(1,0,0)$ does not exist, i.e., the subdominant time scales are dangerous irrelevant variables.

Before discussing some of the phase-transition scenarios allowed for by general reasoning, let us consider some further constraints. For real electronic systems, which interact through long-ranged Coulomb interactions, we have the compressibility sum rule, Eq. (3.133d), which holds for renormalized quantities as well, so that in the long-wavelength limit

$$h+k_s=0 \tag{4.14a}$$

or

$$\gamma_s=-1\ . \tag{4.14b}$$

With Eqs. (4.14) and (4.12c) we see that we have to be careful in defining the dynamical scaling exponent z_3. $\partial n/\partial\mu$, and therefore the screening wave number κ_d, Eq. (2.45b), is not critical at the metal-insulator transition. If we scale the wave numbers in Eq. (3.133d) with ξ, we see that effectively $h+k_s\sim\xi^{1-d}$. We can use this in Eq. (4.12c) to obtain an exact value[15] for z_3,

$$z_3=1\quad\text{(Coulomb interaction)}\ . \tag{4.15a}$$

[15] $z_3=1$ is exact if and only if the charge that enters the screening wave number in Eq. (3.133d) is either not renormalized or flows to a constant at the transition. In principle one cannot rule out the possibility that the charge might flow to infinity at a metal-insulator transition (Fisher *et al.*, 1990). A one-loop calculation for the universality classes MF and MI showed that the charge is not renormalized to that order (Belitz and Kirkpatrick, 1993). Since the charge flowing to infinity would render Eqs. (4.6d) and (4.16b) invalid, it would be important to know whether this result holds to all orders.

Near $d=2$ one has $z_1,z_2=2+O(\epsilon)$, so according to our previous discussion z_3 represents an irrelevant time scale (which is possibly dangerous). We note that for the boson localization problem similar reasoning gives $z=1$ (Fisher *et al.*, 1990) for bosons interacting through long-ranged forces. This exact result has been used for some nontrivial theoretical predictions; see Sec. IX.D. For the fermionic case the result $z_3=1$ is not so useful because the other z's give the asymptotic scaling properties. Finally, for the short-ranged case $H+K^{(s)}=h+k_s$ [see Eq. (3.129a)], and one has, also exactly,

$$z_3=d \quad \text{(short-ranged interaction)} . \tag{4.15b}$$

General theoretical arguments also give a bound on the correlation length exponent for phase transitions in random systems where the control parameter is the disorder (Chayes *et al.*, 1986),

$$\nu \geq \frac{2}{d} . \tag{4.16a}$$

Equations (4.6d) and (4.16a) imply that

$$s \geq \frac{2}{d}(d-2) , \tag{4.16b}$$

i.e., $s \geq \frac{2}{3}$ in $d=3$. Equation (4.16b) is an important inequality that is expected to hold at all metal-insulator transitions. It is generally valid unless either the dimensionless disorder vanishes at the transition, in which case Wegner scaling is not satisfied, or there are multiple divergent lengths at the transition, which makes the application of Eq. (4.16a) ambiguous. The first possibility has been proposed by Castellani, Kotliar, and Lee (1987). It does not seem to be realized in any of the known universality classes, as will be discussed in Sec. VI. The second possibility has never been proposed in the context of the metal-insulator transition. A third possibility, viz., a renormalization of the charge, has been mentioned above in connection with Eq. (4.15a). Since none of the scenarios for a violation of Eq. (4.16b) could be substantiated, we shall assume in what follows that the inequality holds. In Secs. V and VI we discuss possible interpretations of experiments that appear to give $s<\frac{2}{3}$ in $d=3$.

We now briefly discuss a number of scenarios for the interacting metal-insulator transition. Particular attention will be paid to the dynamical scaling exponents. In order of increasing complexity, the following scenarios are possible:

(i) The interactions are irrelevant at the noninteracting fixed point. For this case all of the interaction amplitudes, k_s, k_t, and k_c, that are present scale to zero and h scales to a finite constant. The noninteracting sigma-model field theory is recovered. The dynamical scaling exponents are

$$z_1=z_2=z_3\equiv z=d \tag{4.17a}$$

and

$$y_h=z=d . \tag{4.17b}$$

There are no thermal, magnetic, or density-of-states anomalies at the metal-insulator transition. The interesting physical quantity is the electrical conductivity, which satisfies Eqs. (4.6). This scenario is realized in the case of electrons interacting through short-ranged forces with magnetic impurities or spin-orbit scattering present [universality classes MI(SR) and SO(SR) in Table I]. It will be discussed in Secs. V.A.1 and V.A.3 below.

(ii) The interaction amplitudes and h all scale to a constant (except for k_c, which is absent for this universality class). The deviations of γ_t and γ_s (for the short-ranged case) from their fixed-point values iterate to zero. The latter condition is necessary for a stable RG fixed point that corresponds to a metal-insulator transition with $t=(g^*-g)/g^*$ describing the zero-temperature distance from the transition. Equations (4.12) and (4.15) allow for two distinct subclasses within this scenario, (a) the short-ranged case,

$$z_1=z_2=z_3=d , \tag{4.18a}$$

(b) the long-ranged case,

$$z_1=z_2=d, \quad z_3=1 . \tag{4.18b}$$

There are no thermal or magnetic anomalies at this type of metal-insulator transition. There is a singularity in the single-particle DOS, and so the second nontrivial exponent besides ν is [cf. Eqs. (4.9a) and (4.9b)]

$$\beta=\nu(d-y_h) . \tag{4.19}$$

Notice that in the interacting case the scaling behavior of H does *not* determine the exponent β, since H is not conjugate to the order parameter, and the critical exponent z is independent of y_h.

These cases are realized by systems in an external magnetic field with the Zeeman effect taken into account [universality classes MF(SR) and MF(LR) in Table I]. They will be discussed in Sec. V.A.2.

(iii) The interaction amplitude k_t (or γ_t) is absent (spin diffusion is suppressed by spin-flip or spin-orbit mechanisms) and h scales to zero with the constraint $h+k_s=0$ due to Coulomb interactions. The interaction γ_c may be present, but its fixed-point value is a constant and deviations from γ_c^* are irrelevant. Defining the anomalous dimension of h by κ one has [cf. Eq. (4.12a)],

$$z_1=d+\kappa, \quad z_3=1 . \tag{4.20}$$

There are no magnetic anomalies, but the specific heat satisfies [cf. Eq. (4.11)] $C_v=T\gamma$ with

$$\gamma(t,T)=b^{\kappa}\gamma(b^{1/\nu}t,b^{z_1}T) . \tag{4.21a}$$

Notice that for this case $h(b\to\infty)\to 0$, so $\kappa<0$, and the singular part of γ vanishes at the metal-insulator transition. There is a singularity in the single-particle DOS, so that the singularities at the transition are determined by three unknown exponents: ν, κ, and β.

For this universality class the thermal diffusion coefficient has a different critical behavior than the elec-

trical conductivity or the charge diffusion coefficient given by Eqs. (4.6a) and (4.6b), because of the singularity in C_v. Equations (3.128c), (4.3a), (4.20), and (4.21) yield

$$D_h(t,k,\Omega)=b^{-\varepsilon-\kappa}D_h(b^{1/\nu}t,bk,b^{d+\kappa}T) . \quad (4.21b)$$

This scenario is realized for electrons interacting through long-range forces with magnetic impurities or spin-orbit scattering present [universality classes MI(LR) and SO(LR) in Table I]. It will be discussed in Secs. V.A.1 and V.A.3.

(iv) The interactions scale to infinity, $k_t\to\infty$, $h\to\infty$ and we restrict ourselves to the long-range case, $h+k_s=0$. We also assume that, while h and k_t diverge, γ_t does not, $\gamma_t\to\text{const}<\infty$. There are two dynamical scaling exponents,

$$z_1=z_2=d+\kappa, \quad z_3=1 , \quad (4.22)$$

with $\kappa>0$. There are singularities due to both thermal and magnetic fluctuations as the metal-insulator transition is approached. Since $\gamma\sim h$ and $\chi_s\sim(h+k_t)$ [cf. Eqs. (3.129b) and (3.129c)], these quantities satisfy the scaling equations

$$\gamma(t,T)=b^{\kappa}\gamma(b^{1/\nu}t,b^{d+\kappa}T) , \quad (4.23a)$$

$$\chi(t,T)=b^{\kappa}\chi(b^{1/\nu}t,b^{d+\kappa}T) . \quad (4.23b)$$

In addition there is a singularity in the DOS characterized by the exponent β. The electrical conductivity and the charge diffusion coefficient satisfy Eqs. (4.6) with $z=z_1$. There is also critical spin transport at this phase transition. With Eq. (3.128b) for the spin diffusion coefficient, which is equivalent to $D_s\sim\sigma/\chi_s$, and Eqs. (4.6) and (4.23b) one obtains the scaling equation

$$D_s(t,\mathbf{k},\Omega)=b^{-\varepsilon-\kappa}D_s(b^{1/\nu}t,b\mathbf{k},b^{d+\kappa}\Omega) . \quad (4.24)$$

The thermal diffusion coefficient is similarly given by Eq. (4.21b). The RG flow equations for the critical surface at zero temperature depend on both g and γ_t, and general reasoning allows for two distinct possibilities for a RG stable fixed point describing a metal-insulator transition. In the first scenario $g\to g^*$ and $\gamma_t\to\gamma_t^*$ with the deviations $\delta\gamma_t=\gamma_t-\gamma_t^*$ being irrelevant and the deviations $\delta g=g^*-g$ being relevant and defining the exponent ν. The other, more general, possibility is that some linear combination of g and γ_t defines the relevant distance from the metal-insulator transition, while a different combination gives an irrelevant distance.

This scenario can be realized for the generic universality classes [G(SR) and G(LR) in Table I] for which there is no spin-flip or spin-orbit scattering and no external magnetic fields. It will be discussed in Sec. VI.

Finally, we mention that one could imagine a more complicated scenario for the generic case: $h\to\infty$, $k_t\to\infty$, and $\gamma_t\to\infty$ near the metal-insulator transition. Assuming that $g\to g^*$ at this transition and that the distance $g-g^*$ is relevant, the above possibility implies two relevant perturbations at zero temperature. Such a situation is not consistent with a stable RG fixed point. It is interesting that the $\gamma_t\to\infty$ possibility is realized in a different way, as discussed in the next subsection.

We conclude this subsection by giving in Table II a list of scaling equalities relating the correlation length exponent ν, the conjugate field exponent y_h, and the dynamical scaling exponent z to some of the physical quantities that in principle can be measured near the metal-insulator transition.

B. Possible magnetic phase transitions

In disordered systems electrons diffuse and therefore move more slowly than in clean systems. This implies that any two electrons spend an increased amount of time around each other, which increases the effects of the electron-electron interaction. In general, magnetic phase transitions are triggered by interaction effects. This suggests that in some cases it may be possible for disorder to induce a magnetic phase transition that is distinct from the metal-insulator transition.

To characterize a magnetic phase transition using a nonlinear sigma model designed to describe a metal-insulator transition is not straightforward. For instance, at such a phase transition the physical disorder G is not renormalized. This implies that the dimensionless renormalized disorder g satisfies

$$g(b)\sim b^{-\epsilon}G , \quad (4.25)$$

TABLE II. Physical quantities, their scaling behavior, and scaling equalities relating various exponents.

Physical quantity	Scaling behavior	Scaling equality
Electrical conductivity	$\sigma(t\to0,T=0)\sim t^{s}$	$s=\nu(d-2)$
Density of states	$N_F(t\to0,T=0)\sim t^{\beta}$	$\beta=\nu(d-y_h)$
Order-parameter density susceptibility	$\chi_0(t\to0,T=0)\sim t^{-\gamma}$	$\gamma=\nu(2y_h-d)$
Specific heat	$C_V(t=0,T\to0)\sim T^{1-\kappa/z}$	$\kappa=z-d$
Order-parameter density correlation function	$G(k\to0;t=0,T=0)\sim k^{-2+\eta}$	$\eta=2+d-2y_h$
Heat diffusion coefficient	$D_h(t\to0,T=0)\sim t^{s_h}$	$s_h=s+\nu(z-d)$
Spin diffusion coefficient	$D_s(t\to0,T=0)\sim T^{s_s}$	$s_s=s+\nu\kappa$

i.e., the scaled disorder vanishes as the phase transition is approached ($b \to \infty$). This leaves the scaling behavior of γ_t as the only unknown RG flow equation at zero temperature ($\gamma_s = -1$ for real electronic systems, and we shall see later that γ_c is irrelevant). However, to describe a phase transition with a nonlinear sigma model one should consider the RG flow equation of a coupling constant that has a bare scale dimension of 2. This is because 2 is a lower critical dimension for the model (see the discussion in Sec. III.B.4.a). Equation (4.3a) together with the fact that γ_t is dimensionless motivates defining a bare coupling constant

$$Y = GK^{(t)}/H \tag{4.26a}$$

and the corresponding renormalized coupling constant

$$y = g\gamma_t \ . \tag{4.26b}$$

This suggests that a disorder-induced magnetic phase transition will have the following properties: (i) The dimensionless coupling constant is y given by Eq. (4.26b). If it is a conventional phase transition, y will have a finite value y^* at the fixed point describing the phase transition. The deviation of y from y^*, $t = y^* - y$, is relevant and gives the correlation length exponent. (ii) At the phase transition $g \to 0$ and consequently the dimensionless disorder is an irrelevant variable. $y \to y^*$ ($\neq 0$) and $g \to 0$ implies that at this phase transition $\gamma_t \to \infty$. We conclude that this phase transition is consistent with the scenario that we dismissed at the end of the previous subsection, except that now $g - g^*$ is an irrelevant variable. This implies that, if this fixed point exists, then it is a stable RG fixed point.

Now consider the independent exponents at this phase transition. The zero-temperature distance from the critical point gives the correlation length exponent ν,

$$y^* - y = t(b) = b^{1/\nu} t \ . \tag{4.27a}$$

The specific-heat singularity is related to the dynamical scaling exponent and is given by the anomalous dimension of h,

$$h(b) = b^{\kappa} h \ . \tag{4.27b}$$

Equation (4.25) and $y \to y^*$ yield

$$\gamma_t(b) = b^{\varepsilon} \gamma_t \ , \tag{4.27c}$$

which together with Eq. (4.27b) implies that the triplet interaction amplitude k_t scales as

$$k_t(b) = b^{\varepsilon + \kappa} k_t \ . \tag{4.27d}$$

The relevant susceptibility at this phase transition is the magnetic susceptibility χ_s, which is proportional to k_t (notice that k_t diverges more strongly than h as $b \to \infty$). Equation (4.27d) implies the scaling law

$$\chi_s(t, T) = b^{\varepsilon + \kappa} \chi_s(b^{1/\nu} t, b^z T) \ , \tag{4.28}$$

i.e., the specific-heat exponent gives the susceptibility singularity. To identify the dynamical scaling exponent z, one must use some care. Restricting our considerations to the long-ranged case, we find that $z_3 = 1$ is an irrelevant time scale. The arguments given above and Eqs. (4.12a) and (4.12b) give

$$z_1 = 2 + \kappa \ , \tag{4.29a}$$

$$z_2 = d + \kappa \ . \tag{4.29b}$$

Since $z_2 > z_1$, this naively suggests that τ_1 is an irrelevant time scale and that $z = z_2$. However, in the explicit theory for this phase transition, to be discussed in Sec. VI, the time scale τ_1 acts like a dangerous irrelevant variable, and effectively z in Eq. (4.28) is equal to z_1.

At this magnetic phase transition the electrical conductivity and the mass or charge diffusion coefficient are noncritical. However, the spin diffusion coefficient D_s is singular. With Eqs. (4.28) and (4.25) one obtains

$$D_s(t, \mathbf{k}, \Omega) = b^{-\kappa - \varepsilon} D_s(b^{1/\nu} t, b\mathbf{k}, b^z \Omega) \ . \tag{4.30a}$$

Using Eq. (4.25) and the arguments leading to Eq. (4.21b), one finds that the thermal diffusion coefficient at this phase transition satisfies

$$D_h(t, \mathbf{k}, \Omega) = b^{-\kappa} D_h(b^{1/\nu} t, b\mathbf{k}, b^z \Omega) \ . \tag{4.30b}$$

As already mentioned, in general one expects three independent exponents at zero-temperature phase transitions. So far we have not discussed the scaling of the field conjugate to the order parameter describing the phase transition. Physically we expect that the order parameter is related to the magnetization and that the appropriate susceptibility is just the magnetic susceptibility. The fact that $\chi_s \sim k_t$ and $y = g\gamma_t \to y^*$ at the phase transition gives a constraint that reduces the number of exponents by one. The net result is that there are only two independent exponents at this phase transition. The explicit identification of y_h proceeds as follows. First a source term, S_h, proportional to the magnetic field $\mathbf{h}$ and the magnetization operator $\hat{\mathbf{m}}$, is added to the Grassman action:

$$S_h = \int_0^\beta d\tau \int d\mathbf{x}\, \mathbf{h} \cdot \hat{\mathbf{m}}(\mathbf{x}, \tau) \ . \tag{4.31a}$$

In the Q-matrix field theory S_h is transformed to

$$S_h = -2\pi N_F \sum_{\alpha, n} \int d\mathbf{x} \sum_{r=0,3} \sum_{i=1}^{3} h_i^r Q_{nn}^{\alpha\alpha}(\mathbf{x}) + O(h^2) \ , \tag{4.31b}$$

i.e., the magnetic field couples to the spin-triplet particle-hole components of the Q matrix. Note that, for the magnetic phase transition, the physical field $\mathbf{h}$ couples to a term in the action that is nonlocal in frequency space [see the discussion above Eq. (4.9c)]. Using Eq. (4.7), the fact that the magnetization is given by the first derivative of the free energy with respect to h, and the fact that χ_s is given by the second derivative, one obtains

$$\beta = \nu(z + d - y_h) \ , \tag{4.32a}$$

$$\gamma = \nu(2y_h - z - d) \ . \tag{4.32b}$$

Equations (4.28) and (4.32b) with $z = 2 + \kappa$ give

$$y_h = d + \kappa \ , \tag{4.33a}$$

$$\beta = 2\nu \ . \tag{4.33b}$$

Equation (4.33a) agrees with a result given by Raimondi *et al.* (1990) as a footnote (Ref. 21), which suggests that the dangerous irrelevancy of γ_t might change the result given in the main text of that paper, viz., $y_h = z_1$. Equation (4.33b) naively implies that at the magnetic phase transition the order parameter, viz., the magnetization, will be critical. Things may be more complicated, however. If m_i is the magnetization density in the i direction, then Eqs. (4.7) and (4.31b) give

$$m_i = 2\pi N_F T \sum_n \sum_{r=0,3} \langle {}^r_i Q^{\alpha\alpha}_{nn}(\mathbf{x}) \rangle \ . \tag{4.34}$$

The explicit theory described in Sec. VI suggests that only the component with $n = 0$ is critical. Since this critical component is multiplied by a factor of T in Eq. (4.34), this leads to an uncritical magnetization. According to this interpretation the exponent β given in Eq. (4.33b) is meaningless because the critical quantity has a vanishing prefactor.

V. METAL-INSULATOR TRANSITIONS IN SYSTEMS WITH SPIN-FLIP MECHANISMS

In this section we consider disordered electronic systems that have strong spin-flip or spin-orbit scattering or that are subject to a strong magnetic field. "Strong" in this context means that the spin-flip or spin-orbit scattering rate or the Zeeman splitting is large compared to the temperature. This case, which comprises the universality classes MF, MI, and SO of Table I, turns out to be substantially less complicated than the generic case (class G in Table I), which will be discussed in Sec. VI, because strong magnetic fluctuation effects are suppressed. Unlike the generic case, these systems do not have competing magnetic and metal-insulator transition instabilities. In the classes MF and MI the Cooper channel is also suppressed, which further simplifies the problem. For the metal-insulator transitions in these universality classes a generally accepted theory exists (Finkel'stein, 1983b, 1984a; Castellani, DiCastro, Lee, and Ma, 1984). In the class SO the Cooper channel is present, which makes the problem substantially more complicated. Proposed solutions of the Cooper channel problem (Finkel'stein, 1984b; Castellani, Di Castro, Forgacs, and Sorella, 1984; Kirkpatrick and Belitz, 1993) disagree with each other at a rather fundamental level.

We shall first discuss the asymptotic critical behavior at the metal-insulator transition in systems with magnetic impurities, then in systems with external magnetic fields, and finally in systems with spin-orbit scattering. For the last, we discuss the problems mentioned above in detail. We then turn to the problem of corrections to scaling near the metal-insulator transition. In particular, we discuss a proposal that, for the spin-orbit universality class (and for the generic universality class to be discussed in the next section), there are logarithmic corrections to scaling that make it virtually impossible to probe experimentally the asymptotic critical region in these systems. In Sec. V.B we compare these theoretical results for the metal-insulator transition with experiments on systems that are believed to fall into the respective universality classes.

A. Theory

1. Systems with magnetic impurities

The nonlinear sigma model for this universality class is given by Eqs. (3.118) and (3.92) with the particle-hole spin-singlet restriction

$${}^i_r Q^{\alpha\beta}_{nm} = \delta_{i0}[\delta_{r0}\,{}^0_0 Q^{\alpha\beta}_{nm} + \delta_{r3}\,{}^0_3 Q^{\alpha\beta}_{nm}] \ ; \tag{5.1}$$

i.e., particle-particle and spin-triplet critical fluctuations are suppressed by magnetic impurities. The parameters that appear in this theory, and that will be renormalized by the renormalization procedure, are the disorder G, the frequency coupling constant H, and the spin-singlet particle-hole interaction amplitude $K^{(s)}$.

To proceed, we introduce a number of propagators or, equivalently, vertex functions, which depend on the parameters G, H, and $K^{(s)}$. A loop expansion for each propagator is then developed; see Sec. III.B.3.c. The next step is to introduce renormalized coupling constants and to derive RG flow equations that describe a metal-insulator transition (Castellani, DiCastro, Lee, and Ma, 1984; Finkel'stein, 1984a).

We consider the following (connected) propagators:

$$P^{(1)} = \langle {}^0_0 Q^{\alpha\alpha}_{nn}(\mathbf{x}) \rangle \ , \tag{5.2a}$$

$$P^{(2)}_0 = \langle {}^0_0 q^{\alpha_1\alpha_2}_n(\mathbf{p}_1)\,{}^0_0 q^{\alpha_1\alpha_2}_m(\mathbf{p}_2)^* \rangle_{\alpha_1 \neq \alpha_2} \ , \tag{5.2b}$$

$$P^{(2)}_s = \langle {}^0_0 q^{\alpha\alpha}_n(\mathbf{p}_1)\,{}^0_0 q^{\alpha\alpha}_m(\mathbf{p}_2)^* \rangle \ , \tag{5.2c}$$

and the corresponding vertex functions, $\Gamma^{(1)}$, $\Gamma^{(2)}_0$, and $\Gamma^{(2)}_s$. Here the averages are taken with the full action, and

$${}^0_0 q^{\alpha\beta}_n = (8/n)^{1/2} \sum_{m=1}^{n} {}^0_0 q^{\alpha\beta}_{n-m,-m} \ . \tag{5.2d}$$

For a one-loop calculation of these propagators, it is sufficient to retain terms up to S_4 in Eq. (3.119). The physical meaning of $P^{(1)}$ is that of the one-particle density of states (DOS) at an energy Ω_n away from the Fermi energy [see Eq. (3.131)]. $P^{(2)}_0$ is the basic diffusion propagator and $P^{(2)}_s$ is the charge-density (i.e., singlet) propagator [see Eqs. (3.122)].

To one-loop order $\Gamma^{(1)}$, or $(P^{(1)})^{-1}$, can be obtained by using Eqs. (3.117) to $O(q^2)$ in Eq. (5.2a). Let us first consider the case of short-ranged interactions. With Eqs.

(3.122) one finds

$$\Gamma^{(1)}=1+\frac{\bar{G}}{8\varepsilon}L_s+O(\varepsilon^0\bar{G},\bar{G}^2)\ , \tag{5.3}$$

where we have used dimensional regularization with $\varepsilon=d-2$. In Eq. (5.3),

$$\bar{G}=S_dG/(2\pi)^d\ , \tag{5.4a}$$

$$L_s=\ln(1+K^{(s)}/H)\ , \tag{5.4b}$$

where S_d is the surface area of the d-dimensional unit sphere. Note that for systems interacting through long-ranged interactions $K^{(s)}/H=-1$, and consequently L_s in Eq. (5.3) does not exist. To treat this case we consider the integral, $I_d(\Omega_n)$, leading to the second term in Eq. (5.3). For systems interacting through long-ranged interactions,

$$I_d(\Omega_n)=\frac{G}{8}\sum_{m=1}^{\infty}\sum_{\mathbf{p}}\frac{\Delta\mathcal{D}^s_{n+m}(\mathbf{p})}{n+m}=\bar{G}\frac{GH}{8}\kappa_d^{d-1}\int_0^\infty d\omega\int_0^\infty dp\frac{1}{[\kappa_d^{d-1}p^{3-d}+(\omega+\Omega_n)G\tilde{H}][p^2+(\omega+\Omega_n)GH]}\ . \tag{5.5a}$$

In giving the second equality in Eq. (5.5a) we have used Eqs. (3.134) and neglected terms that are irrelevant for the critical behavior of the density of states, or $\Gamma^{(1)}$. We have also replaced the frequency sum by the appropriate integral in the limit $T\to 0$. First performing the frequency integral and then doing the wave-number integral by dimensionless regularization gives

$$\Gamma^{(1)}=1-\frac{\bar{G}}{4\varepsilon^2}+O(\bar{G}/\varepsilon,\bar{G}^2)\ . \tag{5.5b}$$

Note that the ε^{-2} singularity in Eq. (5.5b) is much stronger than the ε^{-1} singularity usually encountered in expansions about a lower (or upper) critical dimension. Among other things, it implies that the RG flow equation for the single-particle DOS in $d=2$ depends explicitly on the RG scale, and in $d=2+\varepsilon$ the right-hand side is proportional to ε^{-1}. We discuss these features further below.

To one-loop order, there are two topologically distinct diagrams that contribute to the $\Gamma^{(2)}$. They are shown in Fig. 16. Tedious but straightforward calculations give

$$\Gamma_0^{(2)}(\mathbf{k},\Omega_m)=\frac{\mathbf{k}^2}{G}+H\Omega_m+\delta\Gamma_s\ , \tag{5.6a}$$

$$\Gamma_s^{(2)}(\mathbf{k},\Omega_m)=\frac{\mathbf{k}^2}{G}+(H+K^{(s)})\Omega_m+\delta\Gamma_s+\frac{\bar{G}}{8\varepsilon}K^{(s)}\Omega_m(1+2L_s)\ , \tag{5.6b}$$

with

$$\delta\Gamma_s=\frac{(\mathbf{k}^2/G)}{4\varepsilon}\bar{G}\left[1-\frac{H}{K^{(s)}}L_s\right]+\Omega_m\frac{\bar{G}H}{4\varepsilon}\left[L_s-\frac{K^{(s)}}{2H}\right]\ . \tag{5.6c}$$

Note that the factor $\bar{G}/G$ in the first term in Eq. (5.6c) serves only to absorb the factor $S_d/(2\pi)^d$ arising from the momentum integration.

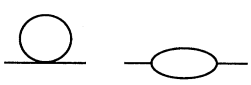

FIG. 16. One-loop diagrams for the two-point vertex functions $\Gamma^{(2)}$.

The characteristic logarithmic structure denoted by L_s in all of these functions arises from performing the summation over frequencies. One way to do the integrals is to perform the wave-number integration first, and then to express the sums over the Matsubara frequencies in terms of Riemann zeta functions. Alternatively, one can transform the Matsubara frequency sums to real frequency integrals using standard methods. Again, the L_s terms in Eq. (5.6) do not exist for systems interacting through long-ranged interactions. However, in the RG flow equations for the coupling constants these terms are canceled by a wave-function renormalization effect. The singularity $L_s\to-\infty$ is apparently relevant only for the DOS.

We next absorb the $\varepsilon\to 0$ singularities encountered in the theory into renormalization constants. The ultimate goal of this procedure is to develop a renormalized field theory with finite coupling constants at the metal-insulator transition. We define a renormalized disorder coupling constant g, a renormalized frequency coupling constant h, and a renormalized interaction amplitude k_s by

$$\bar{G}=\mu^{-\varepsilon}\mathbb{Z}_g g\ , \tag{5.7a}$$

$$H=\mathbb{Z}_h h\ , \tag{5.7b}$$

$$K^{(s)}=\mathbb{Z}_s k_s\ , \tag{5.7c}$$

where μ is an arbitrary momentum scale. The renormalization statement is

$$\Gamma_R^{(N)}(\mathbf{k},\Omega_m;g,h,k_s,\mu)=\mathbb{Z}^{N/2}\Gamma^{(N)}(\mathbf{k},\Omega_m;G,H,K^{(s)})\ , \tag{5.8}$$

where $\Gamma_R^{(N)}$ is the renormalized vertex function and $\mathbb{Z}$ is the field renormalization constant. The three functions $\Gamma^{(1)}$, $\Gamma_0^{(2)}$, and $\Gamma_s^{(2)}$ are sufficient to determine the four renormalization constants. Using minimal subtraction and Eqs. (5.3)–(5.8), one finds

$$\mathbb{Z}=1-\frac{g}{4\varepsilon}l_s+O(g^2)\ , \tag{5.9a}$$

$$\mathbb{Z}_g=1+\frac{g}{4\varepsilon}[1-(1+1/\gamma_s)l_s]+O(g^2)\ , \tag{5.9b}$$

$$\mathbb{Z}_h = 1 + \frac{g}{8\epsilon}\gamma_s + O(g^2) , \tag{5.9c}$$

$$\mathbb{Z}_s = 1 - \frac{g}{8\epsilon} + O(g^2) , \tag{5.9d}$$

with $l_s = \ln(1+\gamma_s)$ and $\gamma_s = k_s/h$.

The one-loop RG flow equations are derived from Eqs. (5.7) and (5.9) in the usual way. With $b \sim \mu^{-1}$ the RG length rescaling factor, one obtains

$$b\frac{dg}{db} = -\epsilon g + \frac{g^2}{4}[1-(1+1/\gamma_s)l_s] + O(g^3) , \tag{5.10a}$$

$$b\frac{dh}{db} = \frac{g}{8}k_s + O(g^2) , \tag{5.10b}$$

$$b\frac{d}{db}(h+k_s) = 0 , \tag{5.10c}$$

$$b\frac{d}{db}\gamma_s = -\frac{g}{8}\gamma_s(1+\gamma_s) . \tag{5.10d}$$

We emphasize two points.

(1) Equation (5.10c) implies that $h+k_s$ is not renormalized to one-loop order. This result is expected to be exact. For systems interacting through screened Coulomb interactions, the compressibility sum rule discussed in Sec. III.B.3.e fixes $h+k_s=0$; see Eq. (3.133d). For systems with short-ranged interactions, $h+k_s$ can be related to the thermodynamic DOS or compressibility $(\partial n/\partial\mu)$; see Eq. (3.129a). This expression involves the density of electrons far from the Fermi energy and consequently should not be sensitive to diffusion corrections.

(2) For systems with short-ranged interactions, Eq. (5.10d) implies that $\gamma_s \to 0$ under RG iterations. This in turn causes the term of $O(g^2)$ in Eq. (5.10a) to vanish. One concludes that in this case interactions are irrelevant and there is no metal-insulator transition to this order. This case realizes the scaling scenario (i) of Sec. IV.A.2. For the noninteracting model, the term of $O(g^3)$, Eq. (5.10a), is needed to recover a metal-insulator transition; see Eq. (3.150b).

We now consider electronic systems interacting through screened Coulomb interactions. In this case there is a metal-insulator transition at one-loop order. For this universality class we have $\gamma_s = -1$ exactly, and Eqs. (5.10a) and (5.10b) become

$$b\frac{dg}{db} = -\epsilon g + \frac{g^2}{4} + O(g^3) , \tag{5.10e}$$

$$b\frac{dh}{db} = -\frac{g}{8}h + O(g^2) . \tag{5.10f}$$

Equation (5.10e) shows a fixed point at $g^* = 4\epsilon$ describing a metal-insulator transition. The RG relevance of g around g^* gives the correlation or localization length exponent ν. If we denote the deviation of g from g^* by t, $t = g^* - g$, then the linearized RG equation for t is

$$b\frac{dt}{db} = \epsilon t . \tag{5.11a}$$

The solution of Eq. (5.11a) is

$$t(b) = b^{\epsilon}t(b=1) = b^{\epsilon}t = b^{1/\nu}t , \tag{5.11b}$$

with $t \equiv t(b=1)$ the physical distance from the metal-insulator transition at zero temperature and frequency and

$$\nu = \frac{1}{\epsilon}[1+O(\epsilon)] . \tag{5.11c}$$

Similarly, the behavior of h in the vicinity of the fixed point gives the critical frequency or temperature dependence. The linearized flow equation is obtained by replacing g in Eq. (5.10f) by $g^* = 4\epsilon$. This yields

$$h(b) \equiv b^{\kappa}h(b=1) = b^{-\epsilon/2}h(b=1) . \tag{5.12a}$$

Equation (5.12a) defines the critical exponent κ. This implies for the dynamical scaling exponent z [cf. Eq. (4.20)]

$$z = d + \kappa = 2 + \epsilon/2 . \tag{5.12b}$$

The scaling theory for this phase transition was discussed in Sec. IV.A.2, scenario (iii). The field theory gives explicit values for the exponents ν and κ (or z), as shown in Eqs. (5.11c) and (5.12). ν also determines the conductivity exponent,

$$s = 1 + O(\epsilon) , \tag{5.12c}$$

and for $T \to 0$ the singular part of the linear specific-heat coefficient behaves as

$$\gamma(t=0, T\to 0) \sim T^{-\kappa/z} . \tag{5.12d}$$

The remaining unknown independent critical exponent is the DOS exponent β defined by Eqs. (4.9b) and (4.19). This can be obtained by studying the scaling behavior of the one-point propagator or vertex function; see Eq. (3.131). To illustrate the points discussed below Eq. (5.5b), we use a normalization-point RG procedure (see, for example, Zinn-Justin, 1989). The normalization condition is

$$\Gamma_R^{(1)}(\Omega_n; g, h)\big|_{gh\Omega_n = \mu^d} = 1 . \tag{5.13a}$$

With Eqs. (5.5b), (5.8), and (5.13a) the wave-function renormalization constant is given to lowest nontrivial order in the disorder by

$$Z^{1/2} = 1 + \frac{g}{8}\kappa_R^{\epsilon}\int_0^{\infty} dx \frac{(1+F_0^s)}{x^2 - (1+F_0^s)x^{3-d}} \times \ln\left\{\frac{\kappa_R^2(1+F_0^s)x^{3-d}+1}{\kappa_R^2 x^2 + 1}\right\} , \tag{5.13b}$$

with $\kappa_R = \kappa_d/\mu \sim b\kappa_d$ the (trivially) renormalized screening wave number. To lowest order in the disorder, the RG flow equation for the scale-dependent vertex function $\Gamma_R^{(1)}$ satisfies

$$\frac{d\ln\Gamma_R^{(1)}}{d\ln b} = \frac{g}{8}J(\kappa_R) , \tag{5.13c}$$

with

$$J(\kappa_R)=2\kappa_R^{2+\varepsilon}\int_0^\infty dx\frac{1+F_0^s}{[1+\kappa_R^2x^2][1+\kappa_R^2(1+F_0^s)x^{3-d}]}\ . \tag{5.13d}$$

The physical critical limit is $\kappa_R\to\infty$ for fixed ε. In this limit,

$$J(\kappa_R)=\begin{cases}\pi/\sin(\pi\varepsilon/2), & d>2,\\ 2\ln\kappa_R, & d=2.\end{cases} \tag{5.13e}$$

Note that incorrect results are obtained by evaluating $J(\kappa_R)$ for $\varepsilon\to 0$ and using the result for $\varepsilon>0$ (Castellani, DiCastro, Lee, and Ma, 1984). With $g^*=4\varepsilon$, Eq. (5.11c), and Eq. (5.13e) in Eq. (5.13c), one obtains (Finkel'stein, 1984a)

$$\beta=1/\varepsilon+O(1)\ . \tag{5.14a}$$

For example, to one-loop order the DOS at the Fermi surface, $N_F(t)=N(t,\Omega=0)$, vanishes as

$$N_F(t)\sim t^{1/\varepsilon}\ . \tag{5.14b}$$

Note that the critical exponent defined by Eq. (5.14b) does not exist in two dimensions ($\varepsilon=0$). Using Eq. (5.13e) and $\ln\kappa_R\sim\ln b$ suggests that $N_F(t)$ vanishes faster than any power law as $t\to 0$ in 2-d systems. The true asymptotic behavior in two dimensions is not known. The explicit scale dependence on the right-hand side of Eq. (5.13c) in two dimensions is due to the fact that the dynamical time scale represented by the exponent z_3 (which is equal to unity) discussed in Sec. IV.A.2 is a dangerous irrelevant variable in this dimension.

Let us discuss this last point in more detail. Conventional momentum-shell renormalization does not allow for an explicit scale dependence in RG flow equations, while in the approach used above it occurs naturally. A connection between these two approaches can be seen if one considers the following points: (1) The screening length is a microscopic length scale, and one expects the critical limit to be $\kappa_d\to\infty$. (2) If κ_d^{-1} were a conventional irrelevant length scale, then the corrections would be of order μ^2/κ_d^2. (3) Such terms always occur in the field-theoretic RG approach if a finite cutoff is used. It is known (Symanzik, 1983) that this scale dependence is equivalent to retaining irrelevant operators in the momentum-shell approach. (4) All of the above arguments apply in the present case, except that here κ_d is related to a dangerous irrelevant operator. This causes $J(\kappa_R=\kappa_d/\mu)$ to diverge in $d=2$ in the limit $\kappa_R\to\infty$. For $d>2$, $J(\kappa_R\to\infty)$ is finite, and the scale dependence of the flow equations vanishes in the critical limit. A detailed technical analysis of the DOS renormalization has been given by Belitz and Kirkpatrick (1993).

The critical behavior of the ultrasonic attenuation for this universality class has been calculated by Dobrosavljevic *et al.* (1991). The result is Eq. (3.151a) with z given by Eq. (5.12b), and

$$x_+=\tfrac{3}{8}\varepsilon^2+O(\varepsilon^3)\ , \tag{5.15a}$$

$$x_-=-1+O(\varepsilon)\ , \tag{5.15b}$$

for the long-range case. For the short-range case one recovers $z=d$ and Eq. (3.152) for the exponents x_+ and x_-.

In Table III we give a summary of the critical exponents for the different universality classes, including those for magnetic impurities discussed here [classes MI(SR) and MI(LR) of Table I]. In Sec. V.B. we shall discuss experimental results relevant for this universality class.

2. Systems in external magnetic fields

The nonlinear sigma model for this universality class is given by Eqs. (3.118) and (3.92) with the restriction

$${}^i_rQ^{\alpha\beta}_{nm}=[\delta_{i0}\delta_{r0}+\delta_{i0}\delta_{r3}+\delta_{i3}\delta_{r0}+\delta_{i3}\delta_{r3}]\,{}^i_rQ^{\alpha\beta}_{nm}\ , \tag{5.16}$$

TABLE III. Values for the three independent exponents ν, z, and β for the eight universality classes of Table I. Values are given for $d=2+\varepsilon$ dimensions except for class G, where approximate values for $d=3$ based on a two-loop approximation are shown. β for class MF(SR) depends on the nonuniversal quantities γ_i^* and $f^*\equiv f(\gamma_i^*)$; see Eqs. (5.25). The critical behavior for class SO(SR) is not understood, and that for class G(SR) has not been considered.

Exponent \ Universality class	MI		MF		SO		G	
	SR	LR	SR	LR	SR	LR	SR	LR
ν	$\frac{1}{2\varepsilon}-\frac{3}{4}\varepsilon+O(\varepsilon^3)$	$\frac{1}{\varepsilon}+O(1)$	$\frac{1}{\varepsilon}+O(1)$	$\frac{1}{\varepsilon}+O(1)$	?	$\frac{1}{\varepsilon}+O(1)$	?	$\cong 0.75$
z	d	$2+\frac{\varepsilon}{2}+O(\varepsilon^2)$	d	d	d	$2+O(\varepsilon^2)$	?	$\cong 5.91$
β	0	$\frac{1}{\varepsilon}+O(1)$	$-(1/2f^*)\times\ln[1-(\gamma_i^*)^2]+O(\varepsilon)$	$\frac{1/2\varepsilon}{1-\ln 2}+O(1)$	0	$\frac{2}{\varepsilon}+O(1)$	?	$\cong 0.50$

i.e., the particle-particle channel and those sectors of the particle-hole channel with a nonvanishing z component of the spin are suppressed by the magnetic field. The parameters that appear in this field theory and that will be renormalized by the renormalization procedure are G, H, $K^{(s)}$, and $K^{(t)}$, with $K^{(t)}$ the spin-triplet particle-hole interaction amplitude.

To proceed, we consider the propagators given by Eqs. (5.2) and, in addition, we introduce the particle-hole spin-triplet propagator,

$$P_t^{(2)}=\langle {}^0_3q_n^{\alpha\alpha}(\mathbf{p}_1){}^0_3q_m^{\alpha\alpha}(\mathbf{p}_2)^*\rangle \ , \tag{5.17}$$

and the corresponding vertex function $\Gamma_t^{(2)}$.

To one-loop order $\Gamma^{(1)}$ can be computed as for the magnetic impurity universality class. Using dimensional regularization with $\varepsilon=d-2$ one finds

$$\Gamma^{(1)}=1+\frac{\bar{G}}{8\varepsilon}(L_s+L_t)+O(\varepsilon^0\bar{G},\bar{G}^2) \ , \tag{5.18a}$$

with

$$L_{s,t}=\ln(1+K^{(s,t)}/H) \ . \tag{5.18b}$$

For systems interacting through long-ranged interactions one finds, to leading order in $1/\varepsilon$,

$$\Gamma^{(1)}=1-\frac{\bar{G}}{4\varepsilon^2}+O(\bar{G}/\varepsilon,\bar{G}^2) \ . \tag{5.19}$$

The remarks given below Eq. (5.5b) also hold for this universality class.

The two-point vertex functions $\Gamma^{(2)}_{0,s,t}$ can be calculated as in the previous subsection. One obtains

$$\Gamma_0^{(2)}(\mathbf{k},\Omega_m)=\frac{\mathbf{k}^2}{G}+H\Omega_m+\delta\Gamma_s+\delta\Gamma_t \ , \tag{5.20a}$$

$$\begin{aligned}\Gamma_s^{(2)}(\mathbf{k},\Omega_m)=&\frac{\mathbf{k}^2}{G}+(H+K^{(s)})\Omega_m+\delta\Gamma_s+\delta\Gamma_t\\&+\frac{\bar{G}}{8\varepsilon}(K^{(s)}+K^{(t)})\Omega_m\\&+\frac{\bar{G}}{4\varepsilon}K^{(s)}\Omega_m(L_s+L_t) \ ,\end{aligned} \tag{5.20b}$$

$$\begin{aligned}\Gamma_t^{(2)}(\mathbf{k},\Omega_m)=&\frac{\mathbf{k}^2}{G}+(H+K^{(t)})\Omega_m+\delta\Gamma_s+\delta\Gamma_t\\&+\frac{\bar{G}}{8\varepsilon}(K^{(s)}+K^{(t)})\Omega_m\\&+\frac{\bar{G}}{4\varepsilon}K^{(t)}\Omega_m(L_s+L_t) \ ,\end{aligned} \tag{5.20c}$$

with

$$\begin{aligned}\delta\Gamma_{s,t}=&\frac{(\mathbf{k}^2/G)}{4\varepsilon}\bar{G}\left[1-\frac{H}{K^{(s,t)}}L_{s,t}\right]\\&+\Omega_m\frac{\bar{G}H}{4\varepsilon}\left[L_{s,t}-\frac{K^{(s,t)}}{2H}\right] \ .\end{aligned} \tag{5.20d}$$

We next absorb the $\varepsilon\to 0$ singularities encountered in the theory into renormalization constants. We define renormalized coupling constants g, h, and k_s by Eqs. (5.7), and a renormalized triplet interaction amplitude k_t by

$$K^{(t)}=\mathbb{Z}_t k_t \ . \tag{5.21}$$

The renormalization statement is

$$\begin{aligned}&\Gamma_R^{(N)}(\mathbf{k},\Omega_m;g,h,k_s,k_t,\mu)\\&\quad=\mathbb{Z}^{N/2}\Gamma^{(N)}(\mathbf{k},\Omega_m;G,H,K^{(s)},K^{(t)}) \ .\end{aligned} \tag{5.22}$$

The four functions $\Gamma^{(2)}$, $\Gamma^{(2)}_{0,s,t}$ are sufficient to determine the five renormalization constants. With minimal subtraction and Eqs. (5.7), (5.20)–(5.22) one finds

$$\mathbb{Z}=1-\frac{g}{4\varepsilon}(l_s+l_t)+O(g^2) \ , \tag{5.23a}$$

$$\begin{aligned}\mathbb{Z}_g=1+\frac{g}{4\varepsilon}&\left\{[1-(1+1/\gamma_s)l_s]+[1-(1+1/\gamma_t)l_t]\right\}\\&+O(g^2) \ ,\end{aligned} \tag{5.23b}$$

$$\mathbb{Z}_h=1+\frac{g}{8\varepsilon}[\gamma_s+\gamma_t]+O(g^2) \ , \tag{5.23c}$$

$$\mathbb{Z}_s=1-\frac{g}{8\varepsilon}[1+k_t/k_s]+O(g^2) \ , \tag{5.23d}$$

$$\mathbb{Z}_t=1-\frac{g}{8\varepsilon}[1+k_s/k_t]+O(g^2) \ , \tag{5.23e}$$

with $l_{s,t}=\ln(1+\gamma_{s,t})$, $\gamma_{s,t}=k_{s,t}/h$.

The one-loop RG flow equations follow from Eqs. (5.7), (5.21), and (5.23) in the usual way. With $b\sim\mu^{-1}$ the RG length rescaling factor one obtains (Castellani, DiCastro, Lee, and Ma, 1984; Finkel'stein, 1984a)

$$\begin{aligned}b\frac{dg}{db}=&-\varepsilon g+\frac{g^2}{4}\{[1-(1+1/\gamma_s)l_s]+[1-(1+1/\gamma_t)l_t]\}\\&+O(g^3) \ ,\end{aligned} \tag{5.24a}$$

$$b\frac{dh}{db}=\frac{g}{8}[k_s+k_t]+O(g^2) \ , \tag{5.24b}$$

$$b\frac{d}{db}(h+k_s)=0 \ , \tag{5.24c}$$

$$b\frac{d}{db}k_t=-\frac{g}{8}[k_s+k_t]+O(g^2) \ . \tag{5.24d}$$

The first comment below Eq. (5.10d) also applies to Eqs. (5.24). The metal-insulator transition implied by Eqs. (5.24) we first consider for systems interacting through short-ranged forces and then separately for systems interacting through screened Coulomb interactions.

The fixed-point values implied by Eqs. (5.24) for systems with short-ranged interactions are

$$h^*=H+\frac{K^{(s)}}{2}+\frac{K^{(t)}}{2} \ , \tag{5.25a}$$

$$k_s^*=-k_t^*=\tfrac{1}{2}(K^{(s)}-K^{(t)}) \ , \tag{5.25b}$$

$$g^*=\frac{4\varepsilon}{f(\gamma_t^*)} \tag{5.25c}$$

with $\gamma_t^*=k_t^*/h^*$ and

$$f(x)=2-\ln(1-x^2)+\frac{1}{x}\ln\left[\frac{1-x}{1+x}\right] . \tag{5.25d}$$

Stability implies the constraint $\gamma_t^*\in[0,1]$. In this interval $f(\gamma_t^*)$ is monotonic, with $f(0)=0$ and $f(1)=2-2\ln 2$.

Using the arguments given below Eqs. (5.10), we find the following critical exponents for the present universality class [class MF(SR) of Table I],

$$\nu=\frac{1}{\varepsilon}[1+O(\varepsilon)] , \tag{5.26a}$$

$$s=1+O(\varepsilon) , \tag{5.26b}$$

$$\kappa=0 , \tag{5.26c}$$

$$z=d , \tag{5.26d}$$

$$\beta=\frac{1}{2f^*}\ln[1-(\gamma_t^*)^2]^{-1}+O(\varepsilon) , \tag{5.26e}$$

where $f^*\equiv f(\gamma_t^*)$. For the critical behavior of the ultrasonic attenuation (Dobrosavljevic *et al.*, 1991) one finds Eq. (3.151a), with z given by Eq. (5.26d), and

$$x_\pm=(\lambda_\pm^*/f^*)\varepsilon/2+O(\varepsilon^2) , \tag{5.26f}$$

where $\lambda_\pm^*=\lambda_\pm(\gamma_t^*)$ with

$$\lambda_\pm(\gamma)=\ln(1-\gamma^2)\pm\{[\ln(1-\gamma^2)]^2+4+16\gamma^2\} . \tag{5.26g}$$

Notice that the critical exponents β and $x_\pm$ are not universal, since γ_t^* depends on bare parameters; see Eqs. (5.25a) and (5.25b). Furthermore, the structure of the theory suggests that at higher order ν will also be nonuniversal. From a RG point of view, nonuniversal critical exponents are very unusual. To understand how they arise in this case one needs to do a stability analysis about the fixed point given by Eqs. (5.25). The coupled equations for g, γ_t, and γ_s lead to one relevant eigenvalue giving the inverse correlation length exponent, one irrelevant eigenvalue, and a zero eigenvalue for the scaling behavior of $\gamma_t-\gamma_s$. The latter is responsible for the nonuniversal critical behavior. This fixed-point scenario is realized only if the zero eigenvalue remains zero to all orders in the loop expansion. This point has so far not been investigated.

For systems interacting through long-ranged Coulomb interactions the fixed-point values implied by Eqs. (5.24) are

$$h^*=-k_s^*=k_t^*=\tfrac{1}{2}(H+K^{(t)}) , \tag{5.27a}$$

$$g^*=\frac{4\varepsilon}{f(1)} , \tag{5.27b}$$

with

$$f(1)=2-2\ln 2 . \tag{5.27c}$$

The critical exponents ν, s, κ, and z are again given by Eqs. (5.26a)–(5.26d). However, the critical exponent β for the DOS is now universal. Equations (5.19) and (5.27) give (Finkel'stein, 1984a)

$$\beta=\frac{1}{\varepsilon f(1)}+O(1) . \tag{5.28a}$$

The critical exponents for the sound attenuation are also universal (Dobrosavljevic *et al.*, 1991),

$$x_+=\frac{5}{4(1-\ln 2)}\varepsilon^2+O(\varepsilon^3) , \tag{5.28b}$$

$$x_-=\frac{-1}{1-\ln 2}+O(\varepsilon) . \tag{5.28c}$$

The scaling theory for these phase transitions has been discussed in Sec. IV.A.2, scenario (ii). In Table III we give a summary of the critical exponents for the different universality classes, including those for systems in magnetic fields [classes MF(SR) and MF(LR) of Table I] discussed above. In Sec. V.B we shall discuss experimental results relevant for these universality classes.

3. Systems with spin-orbit scattering

For the case of spin-orbit scattering, the model is given by Eqs. (3.118) and (3.92), with the restriction

$${}^i_r Q^{\alpha\beta}_{nm}=\delta_{i0}\,{}^0_r Q^{\alpha\beta}_{nm} , \tag{5.29}$$

i.e., all spin-triplet contributions are suppressed by spin-orbit scattering. The parameters that appear in this case are G, H, $K^{(s)}$, and $K^{c(s)}$, where $K^{c(s)}$ is the spin-singlet particle-particle (or Cooperon) interaction amplitude. This case was first discussed by Castellani, DiCastro, Forgacs, and Sorella (1984).

a. Asymptotic critical behavior

As in previous subsections we first determine the asymptotic critical behavior at the metal-insulator transition. We consider the propagators given by Eqs. (5.2). In addition, we need a method for renormalizing $K^{c(s)}$. One way to do this is by directly renormalizing the generating functional (Wilson and Kogut, 1974) rather than the two-point Cooperon propagator (Wilson and Kogut, 1974). This route was taken by Kirkpatrick and Belitz (1991; Belitz and Kirkpatrick, 1992). A crucial point is that $f_n(\mathbf{p})$ in Eqs. (3.122) is logarithmically singular. The asymptotic critical behavior at the metal-insulator transition can be obtained by letting $f_n(\mathbf{p})\to\infty$ at the end of the calculations, thereby neglecting terms that are logarithmically small. We must not, however, neglect the second term in curly brackets in Eq. (3.122c) immediately, since the calculation will produce additional factors of $f_n(\mathbf{p})$ that will appear in the numerator. In the next subsection we shall examine these logarithmically small terms more carefully.

We note at this point that both the approach and the results presented here have been subject to some debate, and that the original treatment of the universality classes

that include Cooperons was different (Castellani, Di Castro, Forgacs, and Sorella, 1984; Finkel'stein, 1984b). Since the full importance of this disagreement becomes apparent only if one goes beyond the leading critical behavior we defer a complete discussion of it to the next subsection. In the present subsection we mention the differences where they occur, but they have no major implications.

To one-loop order, $\Gamma^{(1)}$ can be computed in the same way as for the magnetic impurity universality class. Using dimensional regularization with $\epsilon=d-2$ and the techniques discussed above, one finds

$$\Gamma^{(1)}=1+\frac{\bar{G}}{8\epsilon}L_s+O(\epsilon^0\bar{G},\bar{G}^2)\ . \tag{5.30}$$

For systems interacting through long-ranged interactions one finds, to leading order in $1/\epsilon$,

$$\Gamma^{(1)}=1-\frac{\bar{G}}{4\epsilon^2}+O(\bar{G}/\epsilon,\bar{G}^2)\ . \tag{5.31}$$

The two-point vertex functions $\Gamma^{(2)}_{0,s}$ can be computed as in the previous two subsections. One obtains

$$\Gamma^{(2)}_0(\mathbf{k},\Omega_m)=\frac{\mathbf{k}^2}{G}+H\Omega_m+\delta\Gamma_s-\frac{\mathbf{k}^2/G}{8\epsilon}\bar{G}\ , \tag{5.32a}$$

$$\Gamma^{(2)}_s(\mathbf{k},\Omega_m)=\frac{\mathbf{k}^2}{G}+(H+K^{(s)})\Omega_m+\delta\Gamma_s-\frac{\mathbf{k}^2/G}{8\epsilon}\bar{G}+\frac{\bar{G}}{8\epsilon}K^{(s)}(1+2L_s)\Omega_m\ , \tag{5.32b}$$

with $\delta\Gamma_s$ given by Eq. (5.6c). Notice that the interaction-dependent parts of Eqs. (5.32) are identical with Eqs. (5.6), i.e., in the limit $f_n(\mathbf{p})\to\infty$ the particle-particle interaction amplitude does not contribute to the renormalization of G and H (Finkel'stein, 1984b). The additional contribution compared to Eqs. (5.6) is due to the antilocalization effect of the Cooperon term in the noninteracting part of the action. To compute the $K^{c(s)}$ renormalization, we use the fact that to one-loop order Eq. (3.120c) must have the same form as it does to zero-loop order, except that G, H, and $K^{c(s)}$ acquire one-loop corrections.[16] At zero wave number, the result for $K^{c(s)}$ to this order is

$$K^{c(s),(1\text{-loop})}_{n_1n_2n_3n_4}=K^{c(s)}-\frac{G}{8}K^{(s)}\int_{\mathbf{p}}\{\mathcal{D}^s_{n_3-n_2}(\mathbf{p})+\mathcal{D}_{n_3-n_2}(\mathbf{p})[1-GK^{(s)}|\Omega_{n_2}-\Omega_{n_4}|\mathcal{D}^s_{|n_2-n_4|}(\mathbf{p})]\}\ . \tag{5.33a}$$

Doing the integrals we find

$$K^{c(s)}_{(1\text{-loop})}=K^{c(s)}+\frac{\bar{G}K^{(s)}}{4\epsilon}+O(\epsilon^0\bar{G},\bar{G}^2)\ . \tag{5.33b}$$

Now we again absorb the $\epsilon\to0$ singularities into renormalization constants. We define renormalized coupling constants g, h, and k_s by Eqs. (5.7) and a renormalized spin-singlet particle-particle interaction amplitude $k_{c,s}$ by

$$K^{c(s)}=\mathbb{Z}_{c,s}k_{c,s}\ . \tag{5.34}$$

The renormalization statement is

$$\Gamma^{(N)}_R(\mathbf{k},\Omega_m;g,h,k_s,k_{c,s},\mu)=\mathbb{Z}^{N/2}\Gamma^{(N)}(\mathbf{k},\Omega_m;G,H,K^{(s)},K^{c(s)})\ . \tag{5.35a}$$

Alternatively, one can use renormalized coupling constants and introduce a renormalized q field, q_R, by

$$q=\mathbb{Z}^{1/2}q_R\ . \tag{5.35b}$$

The three functions $\Gamma^{(1)}$, $\Gamma^{(2)}_{0,s}$ and Eqs. (5.33b) and Eqs. (5.35) are sufficient to determine the five renormalization constants. Using minimal subtraction, one finds

$$\mathbb{Z}=1-\frac{g}{4\epsilon}l_s+O(g^2)\ , \tag{5.36a}$$

$$\mathbb{Z}_g=1+\frac{g}{4\epsilon}\{[1-(1+1/\gamma_s)l_s]-\tfrac{1}{2}\}+O(g^2)\ , \tag{5.36b}$$

$$\mathbb{Z}_h=1+\frac{g}{8\epsilon}\gamma_s+O(g^2)\ , \tag{5.36c}$$

$$\mathbb{Z}_s=1-\frac{g}{8\epsilon}+O(g^2)\ , \tag{5.36d}$$

$$\mathbb{Z}_{c,s}=1+\frac{g}{4\epsilon}l_s-\frac{g}{4\epsilon}\frac{k_s}{k_{c,s}}+O(g^2)\ , \tag{5.36e}$$

with $l_s=\ln(1+\gamma_s)$, $\gamma_s=k_s/h$.

For the universality classes with Cooperons it is relevant to note that the vertex functions used in Eq. (5.35) are obtained by taking derivatives with respect to q of the generating functional. These functions are actually matrices in Matsubara frequency space. In Sec III.B.3.b we showed that the process of inverting the vertex functions to obtain the propagators leads to the appearance of the function $f_n(\mathbf{p})$, which was discussed at the beginning of this subsection and which itself is logarithmically singular. This singularity is not eliminated by the renormalization process expressed in Eq. (5.35a), and its effect will be considered in the next subsection.

The one-loop RG flow equations follow from Eqs. (5.7), (5.34), and (5.36) in the usual way. With $b\sim\mu^{-1}$ one obtains

$$b\frac{dg}{db}=-\epsilon g+\frac{g^2}{4}\left\{[1-(1+1/\gamma_s)l_s]-\frac{1}{2}\right\}+O(g^3)\ , \tag{5.37a}$$

[16] This is true provided the theory is renormalizable. See the discussion in the next subsection and in Secs. III.B.3.c and X.

$$b\frac{dh}{db}=\frac{g}{8}k_s+O(g^2)\ , \tag{5.37b}$$

$$b\frac{d}{db}(h+k_s)=0\ , \tag{5.37c}$$

$$b\frac{d}{db}k_{c,s}=-\frac{g}{4}k_s+g\frac{l_s}{4}k_{c,s}+O(g^2)\ . \tag{5.37d}$$

The $(-\frac{1}{2})$ term in Eq. (5.37a) is again the antilocalization effect which is characteristic of systems with spin-orbit scattering (see, for example, Bergmann, 1984; Lee and Ramakrishnan, 1985).

For systems interacting through short-ranged interactions, all of the comments below Eqs. (5.10d) for the short-ranged magnetic impurity universality class apply here as well. That is, the interaction amplitudes γ_s and $k_{c,s}$ flow to zero under RG iterations, and consequently the noninteracting spin-orbit scattering universality class is recovered. This is scaling scenario (i) of Sec. IV.A.2. As was mentioned in Sec. III.C.1, the metal-insulator transition for this universality class is not understood.

As in the previous subsections, for systems interacting through screened Coulomb interactions there is a conventional metal-insulator transition at one-loop order. For this universality class one has $\gamma_s=-1$ exactly, and Eqs. (5.37) become

$$b\frac{dg}{db}=-\varepsilon g+\frac{g^2}{8}+O(g^3)\ , \tag{5.38a}$$

$$b\frac{dh}{db}=-\frac{g}{8}h+O(g^2)\ , \tag{5.38b}$$

$$b\frac{d\gamma_c}{db}=\frac{g}{4}-\frac{g\gamma_c}{2}\left[\frac{1}{\varepsilon}-\frac{1}{4}\right]+O(g^2)\ , \tag{5.38c}$$

with $\gamma_c=k_{c,s}/h$. The second term in Eq. (5.38c) is obtained by the replacement $l_s\to-2/\varepsilon$ for systems interacting through long-ranged interactions. Note that Eq. (5.38c) gives $\gamma_c^*\simeq\varepsilon/2$ and that γ_c is irrelevant in the RG sense. Equations (5.38) are the result obtained by Kirkpatrick and Belitz (1993). The original approach taken by Castellani, Di Castro, Forgacs, and Sorella (1984; see also Finkel'stein, 1984b) considered the Cooper propagator, rather than γ_c, the relevant scaling quantity. The flow equations they obtained are structurally different, but to lowest order in ε they lead to the same asymptotic critical behavior as Eqs. (5.38). In the next subsection we shall see, however, that one approach leads to logarithmic corrections to scaling while the other one does not.

Comparing Eqs. (5.38) with Eqs. (5.10), we see that the critical properties of the spin-orbit universality class are identical to those of the magnetic impurity universality class, except that $g_{\rm SO}^*=2g_{\rm MI}^*$. Therefore the critical exponents for the long-ranged spin-orbit universality class are

$$\nu=\frac{1}{\varepsilon}[1+O(\varepsilon)]\ , \tag{5.39a}$$

$$s=1+O(\varepsilon)\ , \tag{5.39b}$$

$$\kappa=-\varepsilon+O(\varepsilon^2)\ , \tag{5.39c}$$

$$z=2+O(\varepsilon^2)\ , \tag{5.39d}$$

$$\beta=2/\varepsilon+O(1)\ . \tag{5.39e}$$

The scaling theory for this phase transition has been discussed in Sec. IV.A.2, scenario (iii). In Table III we give a summary of the critical exponents for the different universality classes, including those discussed above. In Sec. V.B we shall discuss experimental results relevant for these universality classes.

b. Logarithmic corrections to scaling

In the preceding sections we have emphasized that the localization or metal-insulator transition can be described as a conventional continuous phase transition that occurs at zero temperature. In general the asymptotic critical behavior at such a phase transition is characterized by universal critical exponents, scaling laws relating exponents, and scaling functions. As discussed in the introduction to Sec. IV, corrections to scaling are often important for the interpretation of experiments dealing with continuous phase transitions.

The last point is of particular interest for the case of the metal-insulator transition. The most accurate experiments to date have been performed on a bulk doped semiconductor, viz., Si:P. In this case, a simple power-law fit to the data gives a conductivity exponent s [defined by $\sigma(t,T=0)\sim t^s$], of $s=0.51\pm0.05$ (Thomas, Paalanen, and Rosenbaum, 1983). Taken at face value, this result violates the rigorous inequality given by Eq. (4.16b). Less accurate experiments on other 3-d systems have also led to reported values of s close to, or lower than, $\frac{2}{3}$ (Shafarman, Koon, and Castner, 1989; Dai, Zhang, and Sarachik, 1991a, 1991b, 1992). In general it appears that in systems where time-reversal invariance is not broken (i.e., for the universality classes SO and G one can observe $s\le\frac{2}{3}$. In all other systems the observed values of s are close to 1.

Three mechanisms for a violation of the inequality $s\ge\frac{2}{3}$ have been discussed in connection with Eq. (4.16b). However, a much simpler explanation of the observations than to assume that one of these mechanisms is at work is that the experiments are not in the asymptotic critical region and are measuring preasymptotic, effective exponents that do not need to satisfy the mentioned inequality. In order for this to be a viable explanation, theory must explain why there is such a slow approach to asymptotic scaling, and how corrections to scaling in the experimentally accessible regime can lead to effective exponents with the observed values.

Recently Kirkpatrick and Belitz (1993, 1994) have argued that for time-reversal-invariant systems there are logarithmic corrections to scaling in all dimensions. The most important implication of this conjecture is that it is

very difficult, if not impossible, to reach the asymptotic critical scaling regime experimentally for these universality classes, and that the existing experiments measure an effective conductivity exponent $s_{\rm eff}$ that does not need to satisfy any bound. We shall also see that these corrections to scaling can yield an $s_{\rm eff} \approx 0.5$ in the region of t probed by the experiments, even if the asymptotic critical exponent satisfies $s > \frac{2}{3}$. Other implications will be discussed below.

The general idea behind the logarithmic corrections to scaling is as follows. Consider a time-reversal-invariant system. Suppose the ground state of this system in the clean limit is superconducting. This means that there is an effective Cooper channel interaction amplitude Γ_c whose temperature dependence in the ladder approximation is given by $\Gamma_c \sim \gamma_c^0/[1+\gamma_c^0 \ln(\omega_D/T)]$, with $\gamma_c^0 < 0$ a bare Cooper channel interaction amplitude. The singularity in Γ_c at $T \sim \omega_D \exp(-1/|\gamma_c^0|)$, with ω_D the Debye frequency, signals the Cooper instability leading to the superconducting state. Now consider a system that is not superconducting in the clean limit, so that $\gamma_c^0 > 0$. Γ_c will still appear in the low-temperature description of the system, and it will still have the form given above. In a scaling sense, inverse temperature and the RG length rescaling factor b can be used interchangeably. One therefore expects that, provided γ_c scales to a constant at the metal-insulator transition, there will be a scale-dependent particle-particle interaction amplitude $\Gamma_c \sim \gamma_c^*/(1+\gamma_c^* \ln b)$ with γ_c^* the fixed-point value of γ_c, the renormalized Cooper interaction amplitude. At the metal-insulator transition, $b \to \infty$. This implies that on general grounds, and independent of dimensionality d, we expect an interaction amplitude that vanishes logarithmically as the metal-insulator transition is approached, provided that $\gamma_c \to \gamma_c^*$. This in turn will lead to logarithmic corrections to scaling.

An important question is whether the structural argument given above can be confirmed by repeating the RG analysis of the problem discussed in the previous subsection and retaining those irrelevant terms that vanish most slowly at the transition, i.e., the leading corrections for large but finite $f_n(\mathbf{p})$. Technically we proceed by using a normalization-group RG procedure similar to that used in Sec. V.A.1 to describe the single-particle density of states in systems with long-range interactions. The motivation for this is that the physical argument given above suggests that (1) there might be an explicit scale dependence on the right-hand side of the RG flow equations because Γ_c will appear, and (2) an ultraviolet cutoff, given by ω_D in the physical argument, is explicitly present. A RG description that incorporates both of these features is a normalization-point procedure, with ultraviolet cutoffs on all frequency and wave-number integrals. The resulting flow equations are (Kirkpatrick and Belitz, 1993)

$$b\frac{dg}{db} = -\varepsilon g + \frac{g^2}{8}(1-\Gamma_c) + O(g^3) , \tag{5.40a}$$

$$b\frac{dh}{db} = -\frac{gh}{8}(1-\Gamma_c) + O(g^2) , \tag{5.40b}$$

$$b\frac{d\gamma_c}{db} = \frac{g}{4} + g\frac{\gamma_c}{8}\left[-\frac{4}{\varepsilon} + 1 + \Gamma_c\right] + \frac{g\Gamma_c}{\varepsilon} + O(g^2) . \tag{5.40c}$$

In Eq. (5.40), Γ_c is the Cooper propagator at the normalization point (see below). To zeroth order in the disorder,

$$\Gamma_c = \frac{\gamma_c}{1+\gamma_c \ln b} . \tag{5.40d}$$

At this order the structure of Γ_c is compatible with the phenomenological argument for logarithmic corrections to scaling discussed above. However, the crucial question is whether or not this structure will be retained in the renormalization process. This question has not been answered unambiguously so far, since Γ_c is a very complicated object.

In general, Γ_c in Eqs. (5.40) can be determined by inverting the matrix ${}^0_{1,2}M$ given by Eq. (3.120c), for general $K^{c(s)}_{n_1 n_2, n_1' n_2'}$. At zero temperature the inversion process leads to a Bethe-Salpeter equation that is given by

$$\Gamma_c(\omega,\Omega,\omega') + \int_0^{\omega_0} d\omega'' \frac{\gamma_c(\omega,\Omega,\omega'')\Gamma_c(\omega'',\Omega,\omega')}{(2\omega''+\Omega)} = \gamma_c(\omega,\Omega,\omega') , \tag{5.41a}$$

where ω_0 is a high-frequency cutoff and γ_c is the Cooperon interaction amplitude,

$$(\gamma_c)_{n_1 n_2, n_1' n_2'} = \frac{K^{c(s)}_{n_1 n_2, n_1' n_2'}}{\sqrt{H_{n_1 n_2} H_{n_1' n_2'}}} \to \gamma_c(\omega,\Omega,\omega') , \tag{5.41b}$$

with $H_{n_1 n_2}$ the renormalized value of H. For simplicity, Eqs. (5.41) are given at zero momentum. In the Eqs. (5.40), $\Gamma_c(\omega,\Omega,\omega')$ appears at the normalization point,

$$\Gamma_c = \Gamma_c(\omega,\Omega,\omega')\big|_{\omega=\omega'=\Omega/2=\mu^d/gh} . \tag{5.42}$$

The equations (5.40) by themselves are not a closed set of equations because Γ_c appears on the right-hand side. The general and controversial question is the behavior of Γ_c at the metal-insulator transition. One approach that has been taken in the literature amounts to deriving a differential flow equation for Γ_c by using Eqs. (5.41) and (5.42) (Finkel'stein, 1984b; Castellani, Di Castro, Forgacs, and Sorella, 1984). In an approximation where γ_c is given by the right-hand side of Eq. (5.33a), and where an expansion in powers of γ_c is used, a one-loop RG argument leads to the conclusion that Γ_c approaches a fixed-point value of $O(\varepsilon^{1/2})$ at the metal-insulator transition. An important assumption in this approach is that Γ_c is a scaling operator at the transition. Note that, if this result is correct, then there are no logarithmic corrections to scaling, and the presence of Cooperons has no qualitative implications for the metal-insulator transition.

Arguments against this conclusion have been given by Kirkpatrick and Belitz (1993,1994). This paper also expands on Kirkpatrick and Belitz (1993) in showing how to obtain a different result. Equations (5.40c), (5.41), and (5.42) are also consistent with a fixed-point structure in which γ_c is a constant plus irrelevant corrections, and in which Γ_c vanishes logarithmically slowly, in accord with the physical arguments given above Eqs. (5.40).

At the present time it is not clear which of the two approaches, if either, is the correct one. The fact that two different renormalization procedures yield qualitatively different results casts doubt on the renormalizability of the model (see Sec. X for a discussion of this point). Recent work (Kirkpatrick and Belitz 1994) suggests that more than one renormalization constant is needed for the Cooper interaction channel. Further work on this subject is clearly needed.

Let us discuss the consequences of the logarithmic corrections to scaling, if they exist. We focus on the general structure of the equations (5.40) that are independent of any loop expansion and are presumably valid in $d=3$. First, we notice that the conductivity σ is related to the disorder by

$$\sigma = 4S_d b^{-\epsilon}/(2\pi)^d \pi g(b) \tag{5.43}$$

[cf. Eqs. (3.118), (5.4a), and (5.7a)]. To determine the relevant scale factor b, we consider the general structure of the *linearized* RG equation for the dimensionless distance from the fixed point t [cf. Eq. (5.11a)], viz.,

$$b\frac{dt}{db} = \frac{t}{\nu} - A\frac{\gamma_c^*}{1+\gamma_c^* \ln b} + O((\ln b)^{-2}) \; . \tag{5.44a}$$

Here ν is the correlation length exponent, and A is an unknown coefficient that is universal except for a possible dependence on how t is defined. Solving Eq. (5.44a), and defining b by $t(b)=1$ gives, with Eq. (5.43),

$$\sigma(t\to 0, T=0) \cong \sigma_0 t^s \left[1 + \frac{a_1}{\ln(1/t)} + \frac{a_2}{[\ln(1/t)]^2} + \cdots \right] , \tag{5.44b}$$

with $s=\nu(d-2)$ and t in Eq. (5.44b) the physical distance from the critical point. Here σ_0 is an unknown amplitude, a_1 is determined by A in Eq. (5.44a), a_2 is determined both by A and by the coefficient of the $1/(\ln b)^2$ term in that equation, and so on.

In a similar fashion, the structure of $\sigma(t=0, T\to 0)$ can be determined. The temperature scale is given by $1=\hat{T} b^d h(b)$, where $\hat{T}=T/T_0$ with some nonuniversal temperature scale T_0. The general structure of the linearized equation for h is

$$b\frac{d\ln h}{db} = (z-d) + \frac{\mu\gamma_c^*}{1+\gamma_c^* \ln b} + O((\ln b)^{-2}) \; , \tag{5.45a}$$

with z the dynamical exponent and μ a universal constant. With Eq. (5.43) one obtains

$$\sigma(t=0, T\to 0) = \sigma_0 \hat{T}^{\epsilon/z} [\ln(1/\hat{T})]^{\epsilon\mu/z} \times \left[1 - \frac{\epsilon\mu^2}{z} \frac{\ln\ln(1/\hat{T}^{\epsilon/z})}{\ln(1/\hat{T})} + \frac{b_1}{\ln(1/\hat{T})} + \cdots \right] \tag{5.45b}$$

with b_1 a nonuniversal number. Notice that the asymptotic scaling of $\sigma(t=0, T\to 0)$ is determined by the logarithms. The functional forms given by Eqs. (5.44b) and (5.45b) for the conductivity are expected to be generally valid for time-reversal-invariant systems. It is important to note that if time-reversal invariance is broken, e.g., by a magnetic field, the Cooper channel will acquire a mass and the logarithmic transients will be absent. These results will be used in our discussion of experiments, Secs. V.B and VI.B below.

B. Experiments

There have been an enormous number of experimental studies on disordered electronic systems near the metal-insulator transition. Especially useful for the issues this review deals with has been work that concentrated on the behavior of the conductivity and the dielectric constant at low temperatures (on the order of 10–100 mK) as the metal-insulator is approached. More recently, there also has been some work on the low-temperature behavior of the thermodynamic properties near the metal-insulator transition. All experiments indicate that the conductivity vanishes continuously as the transition is approached, and at least as far as the transport properties are concerned, the metal-insulator transition can be viewed as a conventional continuous phase transition. We shall review charge transport measurements in Sec. V.B.1 below. The thermodynamic properties appear somewhat more mysterious and will be discussed in Sec. V.B.2.

The first difficulty one encounters in comparing theory and experiment is that for a given experimental system it is not always obvious which universality class is relevant. For example, an *a priori* determination of the spin-flip or spin-orbit scattering rate in any system is very difficult. Of course, if a disordered system is subject to a strong magnetic field, then the Zeeman splitting universality class is the relevant one, provided that the spin-orbit and spin-flip scattering rates are small. Kaveh and Mott (1987) have suggested that the strength of the spin-orbit scattering is proportional to $(\Delta Z)^4$, where ΔZ is the difference between the atomic numbers of the host and the impurity atoms. This implies that the spin-orbit universality class should be the relevant one if $\Delta Z \gg 1$ and if the spin-flip scattering rate is small compared to the inverse temperature. Note that Eqs. (5.1), (5.16), and (5.29) imply that a system with both a magnetic field and strong spin-orbit scattering present is in the magnetic impurity universality class, even if it contains no magnetic impurities. It also follows that the universality class MF

can only be realized in systems that in zero field belong to the generic universality class G. While the magnetoresistance has been measured in such systems (for instance, in Si:P, Rosenbaum *et al.*, 1983), no systematic investigation of the critical behavior in a fixed magnetic field has been carried out. We therefore have to limit our discussion below to the universality classes MI and SO.

It should be emphasized that in most experiments it has not been possible to get very close, by phase-transition physics standards, to the metal-insulator transition. This is an intrinsic difficulty related to the fact that disorder is the parameter driving the transition: normally each data point requires the preparation of a separate sample (an exception is the stress-tuning technique that has been applied to Si:P. This material is believed to be in the generic universality class and will be discussed in Sec. VI). As a consequence, the widespread practice of identifying an exponent value, obtained by fitting the available data to a power law, with the actual value of the asymptotic critical exponent is often hard to justify. This point has been discussed before in Secs. IV and V.A.3.b. In general, the exponents obtained experimentally in this way should be regarded as effective exponents, and we shall refer to them as such.

The body of available data related to this problem is so large that we cannot possibly discuss all of it. Since this review is mainly concerned with the metal-insulator transition proper, we shall restrict ourselves to a discussion of some representative systems for which there are some transport data available in the region $t \leq 0.1$, with t the dimensionless distance from the metal-insulator transition. We also limit the discussion to experiments in which the system has been driven through the transition by a systematic variation of disorder. Finally, we consider only experiments in which a careful extrapolation of finite-temperature data to $T=0$ has been performed.

1. Transport properties

As was mentioned above, spin-orbit scattering can be avoided only if the difference in atomic number between host and impurity atoms is very small. Therefore most systems for which the metal-insulator transition has been studied in detail are likely to fall into one of the three universality classes MI, MF, and SO. (An important exception is Si:P, which will be discussed in Sec. VI.) An overview of this large body of work, as well as an extensive collection of references, has been given by Thomas (1985; see also Hirsch, Thomanschefsky, and Holcomb, 1988). Unfortunately, in most systems it has not been possible to obtain more than a few data points in the region $t \leq 0.1$.

One of the best studied materials near the metal-insulator transition is Si:B (Dai, Zhang, and Sarachik, 1991a, 1991b, 1992). For this system $\Delta Z = 9$, and for this and other reasons it is known that it has strong spin-orbit scattering. As discussed above, a strong magnetic field will cause Si:B to be in the magnetic impurity universality class. In Fig. 17 the zero-temperature electrical conductivity is shown as a function of boron concentration n for Si:B in a magnetic field $H=7.5$ T. Here the dimensionless distance from the critical point is $t=n/n_c-1$, with $n_c(H=7.5\ \text{T})=4.22\times10^{18}\ \text{cm}^{-3}$. Note that the smallest t is on the order of $t\cong0.02$. Figure 18 shows the conductivity of ten Si:B samples in a fixed magnetic field of 7.5 T plotted against $T^{1/2}$ for temperatures below 0.5 K. For a comparison of these experimental results with the theory of Sec. V.A, several points should be noted: (i) The zero-temperature data in Fig. 17 were obtained by extrapolating the data in Fig. 18 to $T=0$ assuming a $T^{1/2}$ law. In general this cannot be justified. However, the authors have verified that Fig. 17 is essentially unchanged if, for example, a $T^{1/3}$ law is assumed (Sarachik, 1992). (ii) $d\sigma/dT$ does not change sign as a function of boron concentration in Fig. 18. We shall explain below that this is consistent with Eq. (5.10e) for the magnetic impurity universality class, and that in general one would expect a different behavior in the spin-orbit universality class. (iii) For this experiment, the logarithmic corrections to scaling discussed in Sec. V.A.3.b are not relevant because of the high magnetic field. If we assume that no other slow transients exist, then the effective exponents should be a reasonable approximation to the true asymptotic exponents. An estimate of the conductivity exponent $s(=\nu)$ and the dynamical scaling exponent z can thus be obtained from Figs. 17 and 18. The solid line in Fig. 17 is a best fit to the data assuming the functional form $\sigma=\sigma_0[(n/n_c)-1]^s$. The authors found $\sigma_0=171(^{+23}_{-25})\ (\Omega\,\text{cm})^{-1}$, $n_c=4.22(^{+0.12}_{-0.05})\times10^{18}\ \text{cm}^{-3}$, and $s=1.0(^{+0.10}_{-0.20})$. Similarly, Fig. 18 and Eq. (4.6a) with $\Omega\sim T$ and $\varepsilon=1$ suggest that $z\simeq2$. (iv) The one-loop results in Sec. V.A.1, Eqs. (5.12b) and (5.12c), give $s=\nu=1$ and $z=2.5$. The consistency of these results with experi-

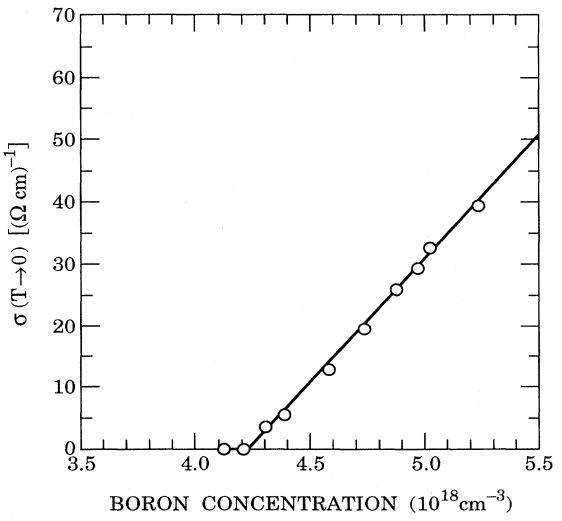

FIG. 17. The zero-temperature conductivity $\sigma(T\to0)$ as a function of dopant concentration in Si:B in a magnetic field $H=7.5$ T. After Dai, Zhang, and Sarachik (1992).

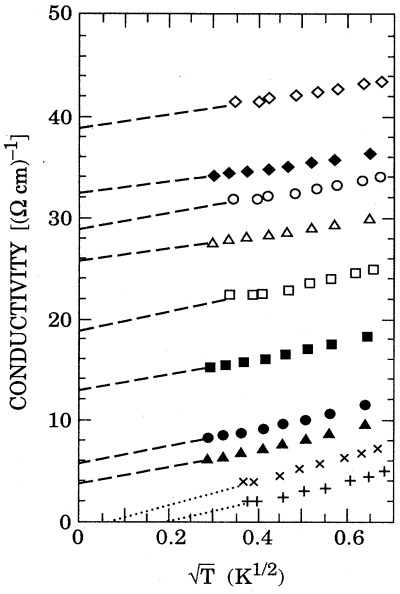

FIG. 18. The conductivity of ten Si:B samples in a magnetic field $H=7.5$ T plotted against $T^{1/2}$ for temperatures below 0.5 K. In units of 10^{18} cm^{-3}, the dopant concentrations are as follows: +, 4.11; ×, 4.2; ▲, 4.30; ●, 4.38; ■, 4.57; □, 4.72; △, 4.86; ○, 4.95; ◆, 5.01; ◇, 5.22. From Dai, Zhang, and Sarachik (1992).

ment is most likely not of great significance, since substantial corrections are to be expected at higher loop order. More relevant is the fact that the measured $s=\nu\geq\frac{2}{3}$, as expected on general theoretical grounds. Finally, in Fig. 19 we test whether or not the experimental results actually satisfy scaling. Except for the sample closest to the metal-insulator transition there is indeed a reasonable collapse of the data onto a single scaling function, even though small breaks in the curve from sample to sample are visible. For the most disordered sample, this break is more substantial. It is currently not known whether the breaks reflect systematic errors in the experiment or whether they indicate a violation of scaling in the available t region.

Si:P in a magnetic field is expected to represent the MF universality class because of the weak spin-orbit scattering in this system. It was recently studied by Dai, Zhang, Bogdanovich, and Sarachik (1993), who found a behavior very similar to that of Si:B in a magnetic field with $s=0.86\pm0.15$. This is remarkable, since Si:P in the absence of a magnetic field shows qualitatively different behavior and displays a value of s ($\cong0.5$) which badly violates the theoretical bound, Eq. (4.16b); see the next section. The fact that a magnetic field restores s to a value larger than $\frac{2}{3}$ is consistent with the interpretation of the zero-field behavior in terms of Cooper-channel-induced logarithmic corrections to scaling; see the last subsection and Sec. VI below.

In Figs. 20 and 21 we show figures analogous to Figs. 17 and 18 for Si:B in the absence of a magnetic field (Dai, Zhang, and Sarachik, 1992). If the spin-flip scattering rate is small then this system should be in the spin-orbit universality class. For this case the logarithmic corrections to scaling discussed in Sec. V.A.3.b can be important. Since in Fig. 20 there are only four data points in the region $t\leq0.1$, the current experimental information about this system is not yet complete enough to allow for a meaningful use of Eq. (5.44b) to fit the data. In Sec. VI.B we shall fit the corrections to scaling to data on Si:P, which are in the generic universality class, and for which more accurate experimental results are available. The solid line in Fig. 20 is a best fit to *all* of the experimental data, covering the region $0.012\leq t\leq0.29$, assum-

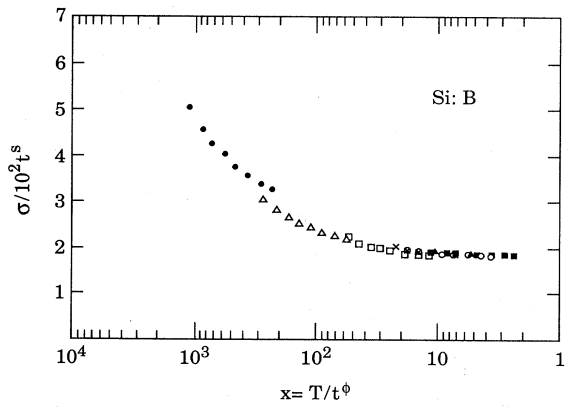

FIG. 19. Scaling plot of the conductivity data shown in Fig. 18. The conductivity exponent is taken to be $s=1$, and the crossover exponent is taken to be $\phi=\nu z=2.0$. The different symbols denote data taken at different disorders. Values of t are as follows: ●, 0.019; △, 0.038; □, 0.083; ×, 0.119; ○, 0.152; ⊗, 0.173; ■, 0.187. The temperature was in the region 0.08 K $<T<$ 0.5 K.

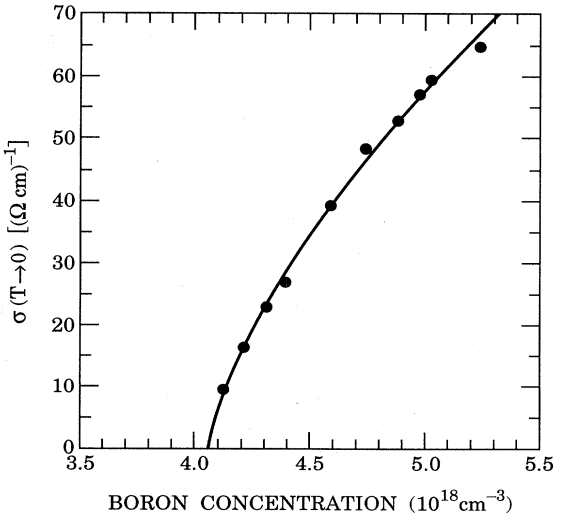

FIG. 20. The zero-temperature conductivity $\sigma(T\rightarrow0)$ as a function of dopant concentration in Si:B in zero magnetic field. After Dai, Zhang, and Sarachik (1992).

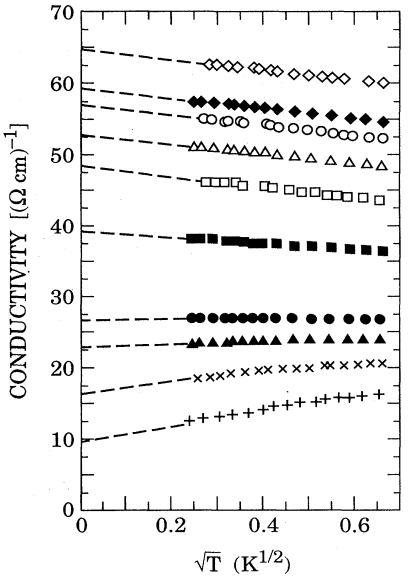

FIG. 21. The conductivity of ten Si:B samples in zero magnetic field plotted against $T^{1/2}$ for temperatures below 0.5 K. The units and symbols are the same as those in Fig. 18. From Dai, Zhang, and Sarachik (1992).

ing the functional form $\sigma=\sigma[(n/n_c)-1]^s$. The fit yielded $\sigma_0=152(^{+10}_{-18})$ $(\Omega\,\mathrm{cm})^{-1}$, $n_c=4.06(^{+0.12}_{-0.02})\times10^{18}$ cm^{-3}, and $s=0.65(^{+0.05}_{-0.14})$. As already emphasized, this value of s should be interpreted as an effective exponent, s_{eff}, and the fact that s_{eff} marginally violates Eq. (4.16b) is of no concern. Note also that Fig. 21 shows that $d\sigma/dT$ changes sign close to the metal-insulator transition. Within the RG description given in Sec. V.A above, this effect can be explained by the change of sign of the coefficient of the g^2 term in Eq. (5.40a). The structure of the one-loop-order term in Eq. (5.40a) allows for such a change of sign to occur. Whether or not $d\sigma/dT$ actually does change sign is a nonuniversal effect that depends on the bare coupling constant γ_c^0 and on the scaling behavior of Γ. Experimentally, it seems that in all systems that show $s_{\mathrm{eff}}<\frac{2}{3}$, $d\sigma/dT$ changes sign close to the metal-insulator transition. Apart from the Si:B system discussed above, this is the case in Si:As (Newman and Holcomb, 1983a), in Si:As,P (Newman and Holcomb, 1983b), and in Si:P (Rosenbaum *et al.*, 1981). The proposed corrections to scaling discussed in Sec. V.A.3.b relate these two features and are consistent with the observations.

Results very similar to those for Si:B described above have been obtained for Si:As (Shafarman, Koon, and Castner, 1989). Nominal zero-temperature data extrapolated from $T\geq0.5$ K in the region $10^{-2}<t<10^{-1}$ yield an effective exponent $s_{\mathrm{eff}}=0.60$. The data on Si:Sb by Long and Pepper (1984) are often also quoted to show $s_{\mathrm{eff}}\cong0.5$. However, this conclusion is based on results far from the transition, while for their data closest to the metal-insulator transition Long and Pepper reported that

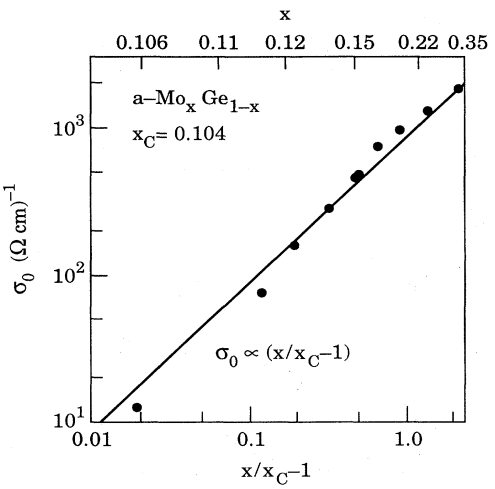

FIG. 22. Extrapolated zero-temperature conductivity σ_0, in a-Mo_xGe_{1-x} vs (x/x_c-1). The critical concentration is $x_c=0.104$. The slope of the straight line gives an effective exponent $s=1.0\pm0.1$. From Yoshizumi *et al.* (1985).

$s_{\mathrm{eff}}\cong1$ yielded a better fit.

In Figs. 22, 23, and 24 we show zero-temperature conductivity data for amorphous Mo_xGe_{1-x} (Yoshizumi, Mael, Geballe, and Green, 1985), amorphous Nb_xSi_{1-x} (Bishop, Spencer, and Dynes, 1985), and amorphous $Si_{1-x}Au_x$ (Nishida *et al.*, 1985). All three of these systems are expected to be in the spin-orbit universality class. They provide an interesting contrast to Si:B. A striking feature of these systems is that the metal-insulator transition (MIT) is very close to a superconductor-metal transition (SMT). For example, in

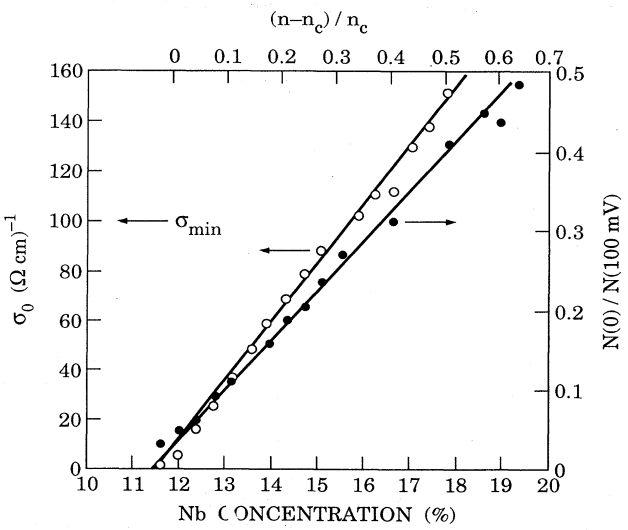

FIG. 23. Extrapolated zero-temperature conductivity σ_0 and tunneling DOS $N(0)$ at $T=2$ K, in amorphous Nb_xSi_{1-x} vs the Nb concentration. The MIT is at a Nb concentration of about 11.5% and the effective critical exponents are $s\cong1$, $\beta\cong1$. After Bishop, Spencer, and Dynes (1985).

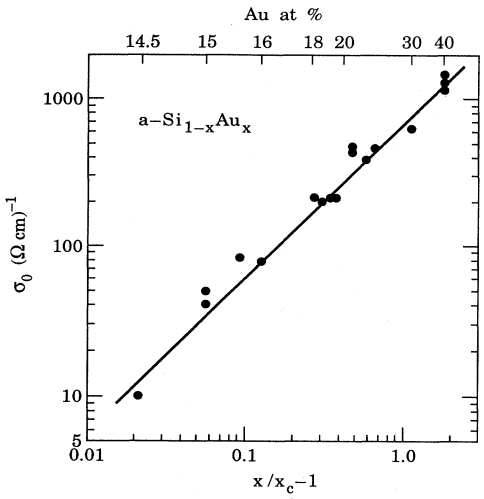

FIG. 24. Extrapolated zero-temperature conductivity σ_0 in amorphous $Si_{1-x}Au_x$ vs (x/x_c-1) with $x_c=0.14$. The effective critical exponent is $s\cong 1$. From Nishida *et al.* (1985).

a-Mo_xGe_{1-x} the MIT occurs at $x\cong 0.10$ and the SMT at $x\cong 0.13$. This suggests that the bare Cooperon interaction amplitude γ_c^0 in these systems near the metal-insulator transition is very small because it must change sign to cause the superconducting instability. This in turn implies that the corrections to scaling discussed in Sec. V.A.3.b might not be important in these systems and that the effective conductivity exponents might be close to the asymptotic ones. These ideas are consistent with the fact that in all three of these systems the effective conductivity exponent is larger than $\frac{2}{3}$ [see Eq. (5.39b)] and with the fact that $d\sigma/dT$ does *not* change sign near the metal-insulator transition in these systems.

The Hall conductivity σ_H has also been measured in many systems that exhibit a metal-insulator transition. In analogy to the longitudinal conductivity, the results have been fitted to a form $\sigma_H=\sigma_1[(n-n_c)/n_c]^{s_H}$. Like s, s_H should be interpreted as an effective exponent. The experimental results can be summarized by saying that systems with $s\cong 1$ show $s_H\cong 1$, while those with $s\lesssim\frac{2}{3}$ show no observable critical behavior of σ_H. Examples of the former behavior are Ge:Sb (Field and Rosenbaum, 1985) and Bi_xKr_{1-x} (Rhode and Micklitz, 1987), while examples of the latter are Si:As (Koon and Castner, 1988) and amorphous Ga_xAr_{1-x} (Zint, Rhode, and Micklitz, 1990). For each of the two cases there is one known exception: AlGaAs:Si has $s=1$ and $s_H=0$ (Katsumoto *et al.*, 1987, 1989), and Si:B has $s\cong 0.5$ and $s_H>0$ (Dai, Zhang, and Sarachik, 1993). Theoretically, the behavior of the Hall conductivity for noninteracting electrons is now understood (see Sec. III.C.2), but the corresponding calculations for the interacting model have not yet been performed and the interpretation of these observations is an open problem.

2. Thermodynamic properties and the tunneling density of states

We now turn to some experimental results for the thermodynamic properties near the metal-insulator transition. Figure 25 shows electron-spin-resonance (ESR) results for the magnetic susceptibility of three samples of Si:P,B (Hirsch *et al.*, 1992). Similar results, albeit at higher temperatures, were obtained earlier for uncompensated Si:B (Sarachik *et al.*, 1985). We note two surprising features. (i) The magnetic susceptibility seems to be singular even in the metallic phase. A power-law fit, $\chi\sim T^{-a}$, gives $a=0.75\pm 0.05$. (ii) The behavior of the magnetic susceptibility is smooth across the transition. This suggests that whatever is causing the magnetic anomalies does not critically depend on what is happening to the conduction electrons. This experimental result suggests that the metallic phase might not be a simple disordered Fermi liquid. None of the theories in Sec. V.A. suggest a magnetic anomaly at the metal-insulator transition, much less in the metallic phase. This singularity in χ is interpreted with very different theoretical ideas in Sec. IX.

Similar conclusions can be reached from measurements of the electronic specific heat of $Si_{1-x}Au_x$ and $C_{1-x}Cu_x$ (LaMadrid, Contrata, and Mochel, 1992), which one expects to be in the spin-orbit universality class. These authors found in the metallic phase, but close to the metal-insulator transition, a temperature dependence of the specific heat $C(T\to 0)\sim T^a$ with $a\cong 0.4$. If one wanted to interpret this as the critical behavior of the specific heat according to Eq. (4.11), one would obtain a dynami-

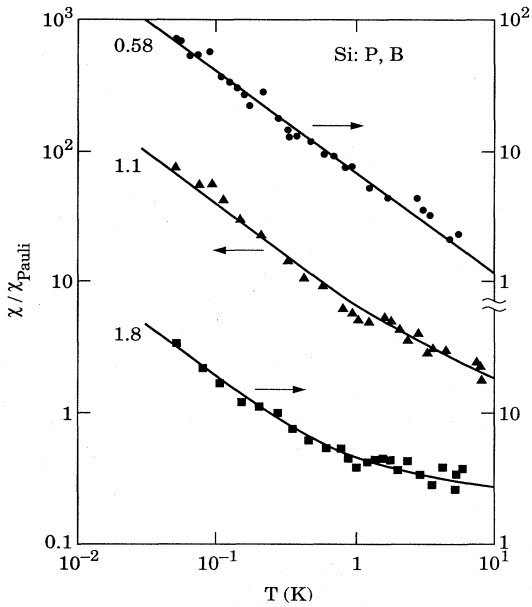

FIG. 25. Temperature dependence of the spin susceptibility χ, normalized to the Pauli value, for three samples of Si:P,B with electron densities $n/n_c=1.8$, 1.1, and 0.58. From Hirsch *et al.* (1992).

cal exponent $z \cong 7.5$. In the light of Eq. (5.39d) it seems unlikely that z is that large. Furthermore, the result is reminiscent of similar measurements on Si:P (Paalanen *et al.*, 1988), which show an anomaly far from, and smooth behavior across, the metal-insulator transition. (We shall discuss this experiment in Sec. VI.) It therefore seems likely that the observed anomaly in the specific heat, like the one in the spin susceptibility, is quite independent of the metal-insulator transition. We shall come back to this in Sec. IX.

We now discuss measurements of the tunneling density of states. The conductance dI/dV of a tunnel junction with a bias voltage V across the junction gives a direct measure of the single-particle or tunneling DOS $N(E)$, with $E = V$ measured from the Fermi surface. Early experiments on $Ge_{1-x}Au_x$ (McMillan and Mochel, 1981) and on granular Al (Dynes and Garno, 1981) showed that $N(0)$, extrapolated to zero temperature, vanishes simultaneously with the electrical conductivity at the metal-insulator transition. For the Nb:Si system discussed in Sec. V.B.1 above this is shown in Fig. 23. The finite value of $N(0)$ at the transition in this case is due to the fact that the temperature dependence of the conductivity was measured down to 10 mK, while the tunneling experiment was done at a fixed temperature $T = 2$ K. The experiment suggests an effective exponent $\beta = 1$. This is certainly consistent with Eq. (5.39e), but no detailed comparison is possible at this point. In any case, all tunneling experiments are consistent with the notion, put forward in Secs. III and IV, that the single-particle DOS is the order parameter for the metal-insulator transition, and that for the Coulomb interaction case the order parameter is critical ($\beta > 0$).

3. Nuclear-spin relaxation

Low-temperature NMR experiments close to the metal-insulator transition have been performed both on the Si nuclei (Paalanen, Ruckenstein, and Thomas, 1985; Hoch and Holcomb, 1988) and on the P nuclei (Alloul and Dellouve, 1987) in Si:P. Note that the magnetic fields applied in these experiments are sufficiently large to put the system in the universality class MF. GeAs and InSb have also been investigated with this technique (Tunstall and Deshmukh, 1979; Tunstall and Sohal, 1983; Tunstall, 1984), and so has compensated Si:(P,B) (Hoch and Holcomb, 1988). The results can be summarized as follows: (1) The spin-lattice relaxation rate $1/T_1$ is strongly enhanced compared to Korringa's (1950) free-electron result and shows anomalous temperature and magnetic field dependences (cf. Figs. 26 and 27). (2) $1/T_1$ and all other NMR observables show a smooth behavior if the system is driven through the metal-insulator transition. In fact, it is impossible to determine the critical concentration for the transition from NMR experiments (Jerome *et al.*, 1985; Paalanen, Ruckenstein, and Thomas, 1985; Alloul and Dellouve, 1987). The

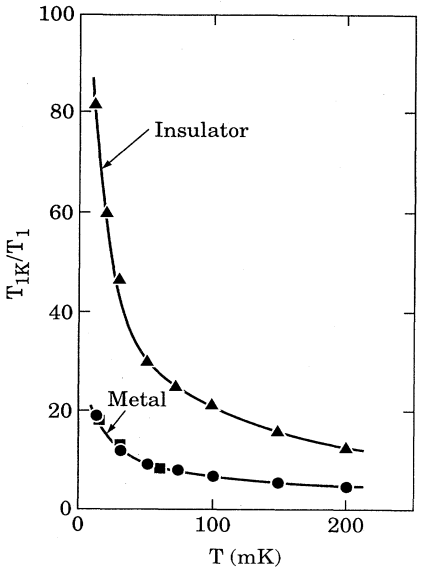

FIG. 26. Temperature dependence of the spin-lattice relaxation time T_1 of ^{29}Si nuclei in Si:P. The data were taken on ▲, an insulating sample ($n/n_c = 0.90$); and two metallic samples—●, $n/n_c = 1.03$; ■, $n/n_c = 1.035$—in a magnetic field $B = 0.844$ T. $1/T_1$ is normalized by the Korringa rate $1/T_{1K}$. Note that $1/T_{1K}$ has a linear temperature dependence. The lines are guides to the eye. From Paalanen, Ruckenstein, and Thomas (1985).

spin-lattice relaxation rate has this feature in common with the spin susceptibility discussed above.

The experimentalists have interpreted these results as evidence for the presence of quasistatic spins or local magnetic moments in the material, or as evidence for spin localization even on the metallic side of the metal-insulator transition. On the theoretical side there have not been many attempts to interpret these observations. Götze and Ketterle (1983) have shown, within a model of noninteracting electrons, that diffusive electron dynamics

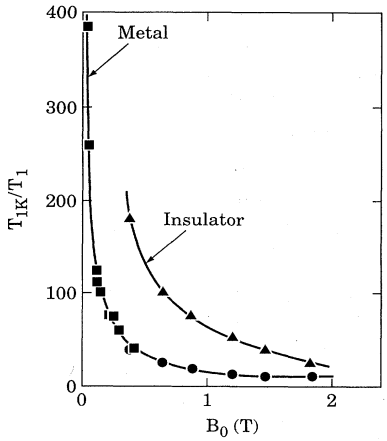

FIG. 27. Same as Fig. 26 as a function of the static magnetic field B_0 at temperature $T = 13.5$ mK. After Paalanen, Ruckenstein, and Thomas (1985).

lead to an enhancement of $1/T_1$ over the Korringa value. Gan and Lee (1986) have interpreted the experiment of Paalanen, Ruckenstein, and Thomas (1985) in terms of local magnetic moments. We shall discuss this work in Sec. IX.B.4 below. For the field theory discussed in Sec. V.A the NMR response has not been calculated so far.

VI. PHASE TRANSITIONS IN THE ABSENCE OF SPIN-FLIP MECHANISMS

In this section we examine the disordered electron problem in the absence of spin-flip or spin-orbit scattering mechanisms. This model is presumed to be relevant in situations where the spin-flip or spin-orbit scattering rate is small compared to the temperature. In some systems these rates are so small that this condition is fulfilled down to temperatures on the order of 1 mK. Such systems realize the generic universality class for all practical purposes. This case is considerably more complicated than the models considered in Sec. V. The Cooper channel is present in this universality class, which causes the same problems as were discussed in Sec. V.A.3. However, even if one ignores this problem as has been done in a large part of the literature, following Finkel'stein (1983a), who neglected the Cooper channel, the particle-hole spin-triplet channel poses a separate problem: the interaction amplitude γ_t flows to infinity under RG iterations for any disorder in $d=2$, and before the metal-insulator transition is reached in $d=2+\varepsilon$. That γ_t poses a runaway problem was recognized early on (Castellani, DiCastro, Lee, Ma, Sorella, and Tabet, 1984; Finkel'stein, 1984b, 1984c).

A considerable body of literature has tried to make physical sense out of the runaway behavior of γ_t based on one-loop calculations. It has been proposed that $\gamma_t \to \infty$, which implies a diverging spin susceptibility according to Eq. (3.129b), somehow signals the formation of local magnetic moments (Castellani, DiCastro, Lee, Ma, Sorella, and Tabet, 1984, 1986; Finkel'stein, 1984c). This suggestion was in part motivated by experiments on Si:P, supposedly in the G universality class, which showed a divergent magnetic susceptibility in the metallic phase that could be interpreted in terms of local moments (Paalanen *et al.*, 1986). To explain this observation within the field theory it was argued that a long-wavelength theory is not really suitable for describing local-moment formation and that it could only be expected to crudely signal a magnetic anomaly. It was speculated that a metal-insulator transition would occur at larger disorder against a background of an inhomogeneous spin density. However, this scenario ignored earlier data on the SO-class material Si:B (Sarachik *et al.*, 1985), which showed a very similar singularity. Allen *et al.* (1993) and Belitz and Kirkpatrick (1994) have stressed that a divergent spin susceptibility in the metallic phase is a very common phenomenon in class-G systems, where γ_t flows to infinity, as well as in others for which there is no theoretical basis for assuming a runaway behavior of γ_t. This rules out the unusual scaling behavior of γ_t in class G as an explanation for the experimental susceptibility anomaly. We shall discuss the spin susceptibility anomaly and theoretical explanations for it in detail in Sec. IX.

A different proposal was made by Castellani, Kotliar, and Lee (1987). These authors noticed that the one-loop RG flow equations did contain a fixed point where γ_t scales to infinity and the renormalized disorder g scales to zero, such that the product $y=g\gamma_t$ is finite. Their suggestion was that this fixed point described an unorthodox metal-insulator transition in which the renormalized disorder vanishes with an exponent ϑ and Wegner scaling is violated; see the discussion of Eqs. (4.6a) and (4.16b) in Sec. IV. Kirkpatrick and Belitz (1989) followed up on this suggestion. They proved that $s=0$ to all orders in the loop expansion, which would be consistent with a discontinuous behavior of the conductivity at the unorthodox metal-insulator transition. However, shortly thereafter it was noticed that logarithmic terms in the γ_t flow equation at two-loop order ruin the fixed-point scenario of Castellani, Kotliar, and Lee (Belitz and Kirkpatrick, 1989b).

The next development was the realization (Kirkpatrick and Belitz, 1990b, Belitz and Kirkpatrick, 1991) that the leading logarithmic terms in the γ_t flow could be controlled to all orders in the limit $\gamma_t \to \infty$. A resummation of the perturbation theory to all orders was performed, which resulted in coupled integral equations for the spin and heat diffusion coefficients. Numerical and asymptotic analytical solutions showed that, for sufficiently large values of the Fermi-liquid parameter F_0^a, i.e., for a sufficiently large spin susceptibility in the clean limit, the equations describe a phase transition from a Fermi liquid to a metallic phase with an infinite spin susceptibility. The runaway behavior of γ_t was thus linked to a magnetic or pseudomagnetic phase transition that is distinct from, and precedes, the metal-insulator transition. The qualitative features of the solution of the integral equations were later confirmed, and the critical behavior studied in more detail, by means of RG analysis (Kirkpatrick and Belitz, 1992b). The nature of the magnetic or pseudomagnetic phase is not entirely clear yet, but it has been speculated that it shows no long-range order but rather a glassy state of the spins. Based on this speculation, this phase has been dubbed the incompletely frozen spin (IFS) phase. Attempts to interpret the spin susceptibility anomaly in Si:P as evidence for the IFS phase (Belitz and Kirkpatrick, 1991) were mistaken for the same reason as the local-moment interpretations of $\gamma_t \to \infty$ mentioned above. The IFS phase has probably not been observed so far. We shall come back to this in Sec. VI.B below.

For values of F_0^a smaller than a critical value, both the integral equation approach and RG analysis suggest that γ_t does not flow to infinity. This has led to the proposal that, in that region of phase space, there is a direct Fermi-liquid-to-insulator (FL-I) phase transition that is intrinsically nonperturbative in nature (Kirkpatrick and

Belitz, 1992a). The resulting phase diagram is shown in Fig. 28 for three different dimensionalities. The explicit description of the FL-I transition is technically on much weaker ground than the FL-IFS phase transition, since for the latter the asymptotic smallness of $1/\gamma_t$ provides the theory with an additional small parameter that allows for an exact treatment of the spin-triplet channel. This advantage is absent if γ_t does not diverge, and the description of the presumed metal-insulator transition is not controlled in the usual small-parameter sense.

In the theoretical part of this section we describe how to derive these results. As in Sec. V, we first examine the one-loop renormalization of the coupling constants and construct the RG flow equations to one-loop order. These equations do not show a fixed point corresponding to a conventional metal-insulator transition. Rather, they display the mentioned runaway RG flow trajectory for the triplet particle-hole interaction amplitude, $\gamma_t = k_t/h$. We then discuss the analysis of the theory in the limit $\gamma_t \to \infty$, which shows that this runaway flow can be interpreted as a phase transition, unrelated to the metal-insulator transition, which is consistent with the magnetic phase transition scenario discussed in Sec. IV.B. The theory of this transition predicts the existence of a lower critical dimension d_c, with $2 < d_c < 3$, above which the FL-IFS phase transition does not extend to $F_0^a = 0$. This suggests the phase diagram shown in Fig. 28. For sufficiently large dimensionality there is a direct FL-I transition at which γ_t stays finite. In Sec. VI.A.3 an approximate theory is given to describe this phase transition in $d = 3$. If the logarithmic corrections to scaling discussed in Sec. V.A.3.b exist, they will be present in the universality class G as well as in class SO, and they may be important at this metal-insulator transition. This topic is further discussed in Sec. VI.B, where we examine some representative experimental data for this universality class.

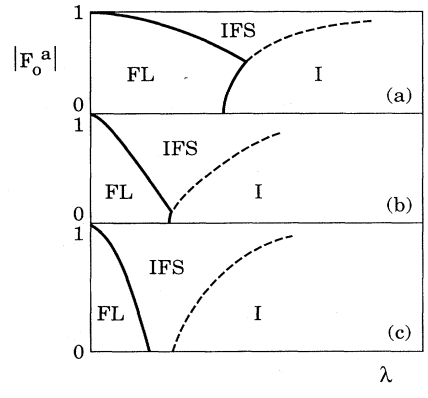

FIG. 28. Qualitative phase diagram for the generic universality class. FL, IFS, and I denote the Fermi liquid, incompletely frozen spin, and insulator phases, respectively. λ is a dimensionless measure of disorder, e.g., $\lambda = Gk_F^{d-2}$. The phase diagram is shown for three different dimensionalities relative to the critical value d_c: (a) $d > d_c$, (b) $d \gtrsim d_c$, (c) $d < d_c$. No information is presently available for the IFS-I transition (dashed lines). From Kirkpatrick and Belitz (1992b).

A. Theory

1. The loop expansion

The nonlinear sigma model for this universality class is given by Eqs. (3.118) and (3.92) with no restrictions on the ${}_i^r Q_{nm}^{\alpha\beta}$. The parameters that appear in the theory and will be renormalized by the renormalization procedure are the disorder G, the frequency coupling constant H, the spin-singlet and triplet particle-hole interaction amplitudes $K^{(s)}$ and $K^{(t)}$, and the spin-singlet particle-particle interaction amplitude $K^{c(s)}$. We shall start with a determination of the one-loop renormalizations of these coupling constants. In contrast to the situation in Sec. V, however, the results of the one-loop calculation will not be conclusive, and we shall be forced to consider the theory to all orders in the loop expansion. This turns out to be possible in the limit $\gamma_t \to \infty$, where $1/\gamma_t$ can be used as a small parameter.

To proceed we consider the propagators given by Eqs. (5.2) and (5.17) and the corresponding vertex functions, $\Gamma^{(1)}$, $\Gamma_0^{(2)}$, $\Gamma_s^{(2)}$, and $\Gamma_t^{(2)}$. To renormalize $K^{c(s)}$ we use the method discussed in Sec. V.A.3. As already mentioned, there may be important logarithmic corrections to scaling for this universality class. This possibility has been discussed in Sec. V.A.3.b, both specifically for the spin-orbit universality class and from a general point of view. In the present subsection we ignore these effects for simplicity. We will take them into account, however, in comparing theory with experiment in Sec. VI.B. Of course, all caveats about the treatment of the Cooper channel that were mentioned in Sec. V apply here as well.

a. One-loop results

To one-loop order $\Gamma^{(1)}$ can be computed as in the previous section. Using dimensional regularization with $\epsilon = d - 2$ and the techniques discussed in Sec. V, one finds

$$\Gamma^{(1)} = 1 + \frac{\bar{G}}{8\epsilon}(L_s + 3L_t) + O(\epsilon^0 \bar{G}, \bar{G}^2) \ . \tag{6.1}$$

Here and in the remainder of this section we use the same notation as in Sec. V. $\bar{G}$ and $L_{s,t}$ are given by Eqs. (5.4a) and (5.18b), respectively. For systems interacting through long-ranged interactions, $L_s \to -\infty$ and the one-loop term in Eq. (6.1) does not exist. For this case the considerations discussed in connection with Eqs. (5.5) give

$$\Gamma^{(1)} = 1 + \frac{\bar{G}}{8\epsilon}\left[3L_t - \frac{2}{\epsilon}\right] + O(\bar{G}/\epsilon, \bar{G}^2) \ . \tag{6.2}$$

The two-point vertex functions $\Gamma_{0,s,t}^{(2)}$ can also be computed as in Sec. V. One obtains

$$\Gamma_0^{(2)}(\mathbf{k},\Omega_m)=\frac{\mathbf{k}^2}{G}+H\Omega_m+\delta\Gamma_s+3\delta\Gamma_t\ , \tag{6.3a}$$

$$\Gamma_{s,t}^{(2)}(\mathbf{k},\Omega_m)=\frac{\mathbf{k}^2}{G}+(H+K^{(s,t)})\Omega_m+\Delta\Gamma_{s,t}\ , \tag{6.3b}$$

with

$$\Delta\Gamma_s(\mathbf{k},\Omega_m)=\delta\Gamma_s+3\delta\Gamma_t+\frac{\bar{G}}{8\varepsilon}\Omega_m(K^{(s)}+3K^{(t)})+\frac{\bar{G}}{4\varepsilon}\Omega_m K^{(s)}(L_s+3L_t)\ , \tag{6.3c}$$

$$\Delta\Gamma_t(\mathbf{k},\Omega_m)=\delta\Gamma_s+3\delta\Gamma_t+\frac{\bar{G}}{8\varepsilon}\Omega_m(K^{(s)}-K^{(t)})-\frac{\bar{G}}{2\varepsilon}\Omega_m H(K^{(t)}/H)^2+\frac{\bar{G}}{4\varepsilon}\Omega_m K^{(t)}(L_s+3L_t)\ , \tag{6.3d}$$

with $\delta\Gamma_{s,t}$ given by Eq. (5.20d). To compute the $K^{c(s)}$ renormalization we use the same method as in Sec. V.A.3.a. To one-loop order we find for $K^{c(s)}$

$$K^{c(s),(1\text{-loop})}_{n_1n_2,n_3n_4}=K^{c(s)}-\frac{G}{8}\int_{\mathrm{p}}\{K^{(s)}\mathcal{D}^s_{n_3-n_2}(\mathbf{p})-3K^{(t)}\mathcal{D}^t_{n_3-n_2}(\mathbf{p})+K^{(s)}\mathcal{D}_{n_3-n_2}(\mathbf{p})[1-GK^{(s)}|\Omega_{n_2}-\Omega_{n_4}|\mathcal{D}^s_{|n_2-n_4|}(\mathbf{p})]$$
$$-3K^{(t)}\mathcal{D}_{n_3-n_2}(\mathbf{p})[1-GK^{(t)}|\Omega_{n_2}-\Omega_{n_4}|\mathcal{D}^t_{|n_2-n_4|}(\mathbf{p})]\}\ , \tag{6.4a}$$

or

$$K^{c(s)}_{(1\text{-loop})}=K^{c(s)}+\frac{\bar{G}}{4\varepsilon}(K^{(s)}-3K^{(t)})+O(\varepsilon^0\bar{G},\bar{G}^2)\ . \tag{6.4b}$$

For later use we also note that at one-loop order there is a disorder-induced triplet particle-particle interaction amplitude that is given by

$$K^{c(t),(1\text{-loop})}_{n_1n_2,n_3n_4}=\frac{\bar{G}^2}{8}\int_{\mathrm{p}}\{-(\Omega_{n_3}-\Omega_{n_2})\mathcal{D}_{n_3-n_2}(\mathbf{p})[(K^{(s)})^2\mathcal{D}^s_{n_3-n_2}(\mathbf{p})+(K^{(t)})^2\mathcal{D}^t_{n_3-n_2}(\mathbf{p})]$$
$$+|\Omega_{n_2}-\Omega_{n_4}|\mathcal{D}_{n_3-n_2}(\mathbf{p})[(K^{(s)})^2\mathcal{D}^s_{|n_2-n_4|}(\mathbf{p})+(K^{(t)})^2\mathcal{D}^t_{|n_2-n_4|}(\mathbf{p})]\}\ . \tag{6.5}$$

There are no terms of order ε^{-1} in Eq. (6.5), and we ignore $K^{c(t)}$ in the remainder of this section. We shall come back to it in Sec. VIII.

We next absorb the $\varepsilon\to 0$ singularities encountered in the theory into renormalization constants. We define renormalized coupling constants g, h, k_t, k_s, and $k_{c,s}$ by Eqs. (5.7), (5.21), and (5.34). The renormalization statement is

$$\Gamma_R^{(N)}(\mathbf{k},\Omega_m;g,h,k_s,k_t,k_{c,s},\mu)=\mathbb{Z}^{N/2}\Gamma^{(N)}(\mathbf{k},\Omega_m;G,H,K^{(s)},K^{(t)},K^{c(s)})\ . \tag{6.6}$$

With minimal subtraction one finds that the renormalization constants are

$$\mathbb{Z}=1-\frac{g}{4\varepsilon}(l_s+3l_t)+O(g^2)\ , \tag{6.7a}$$

$$\mathbb{Z}_g=1+\frac{g}{4\varepsilon}\{[1-(1+1/\gamma_s)l_s]+3[1-(1+1/\gamma_t)l_t]+1\}+O(g^2)\ , \tag{6.7b}$$

$$\mathbb{Z}_h=1+\frac{g}{8\varepsilon}(\gamma_s+3\gamma_t)+O(g^2)\ , \tag{6.7c}$$

$$\mathbb{Z}_s=1-\frac{g}{8\varepsilon}(1+3k_t/k_s)+O(g^2)\ , \tag{6.7d}$$

$$\mathbb{Z}_t=1+\frac{g\gamma_t}{2\varepsilon}+\frac{g}{8\varepsilon}(1-k_s/k_t)+O(g^2)\ . \tag{6.7e}$$

$$\mathbb{Z}_{c,s}=1+\frac{g}{4\varepsilon}(l_s-3l_t)-\frac{g}{4\varepsilon k_{c,s}}(k_s-3k_t)+O(g^2)\ , \tag{6.7f}$$

with $l_{s,t}=\ln(1+\gamma_{s,t})$, $\gamma_{s,t}=k_{s,t}/h$ as in Sec. V.

The one-loop RG flow equations follow from Eqs. (5.7), (5.21), (5.34), and (6.7) in the usual way. With $b\sim\mu^{-1}$ one obtains (Finkel'stein, 1983a; Castellani, DiCastro, Lee, and Ma, 1984; Kirkpatrick and Belitz, 1993)

$$b\frac{dg}{db}=-\varepsilon g+\frac{g^2}{4}\{5-(1+1/\gamma_s)l_s-3(1+1/\gamma_t)l_t\}+O(g^3)\ , \tag{6.8a}$$

$$b\frac{dh}{db}=\frac{g}{8}h[3\gamma_t+\gamma_s]+O(g^2)\ , \tag{6.8b}$$

$$b\frac{d}{db}(h+k_s)=0\ , \tag{6.8c}$$

$$b\frac{d}{db}k_t=\frac{gk_t\gamma_t}{2}+g\frac{k_t}{8}-\frac{g}{8}k_s+O(g^2)\ , \tag{6.8d}$$

$$b\frac{d}{db}k_{c,s}=\frac{g}{4}(3k_t-k_s)+\frac{g}{4}k_{c,s}(l_s+3l_t)+O(g^2)\ . \tag{6.8e}$$

For future use we note

$$b\frac{d\gamma_t}{db}=\frac{g\gamma_t^2}{8}+\frac{g\gamma_t}{8}(1-\gamma_s)-\frac{g\gamma_s}{8}+O(g^2)\ , \tag{6.8f}$$

$$b\frac{d\gamma_c}{db}=\frac{g}{4}(3\gamma_t-\gamma_s)+\frac{g\gamma_c}{4}\left[3l_t+l_s-\frac{3\gamma_t}{2}-\frac{\gamma_s}{2}\right]+O(g^2)\ , \tag{6.8g}$$

with $\gamma_c=k_{c,s}/h$. For systems interacting through long-range forces, l_s in Eqs. (6.8e) and (6.8g) should be replaced by $-2/\varepsilon$; cf. Eqs. (5.5).

We emphasize a few points concerning the structure of these equations.

(1) For the general reasons discussed in Sec. V, $h+k_s$ is not renormalized [see Eq. (6.8c)]. Further, Eqs. (6.8b) and (6.8c) imply that both h and $|k_s|$ increase under RG iterations such that $\gamma_s\to-1$. Therefore for this universality class a short-ranged electron-electron interaction leads, after some initial transient region, to the same flow behavior for the coupling constants as the Coulomb interaction (Castellani, DiCastro, Lee, and Ma, 1984).

(2) An examination of Eqs. (6.8) shows that they do not have a fixed point with $g\to g^*\neq 0$. That is, they do not describe a metal-insulator transition.

(3) Motivated by the general discussion given in Sec. IV.B, we next look for a nontrivial fixed point at $g\to 0$, $\gamma_t\to\infty$ such that $y=g\gamma_t$ is finite (Castellani, Kotliar, and Lee, 1987). In this limit the RG flow equations are

$$b\frac{dg}{db}=-\varepsilon g+O(g^3)\ , \tag{6.9a}$$

$$b\frac{dh}{db}=\frac{3}{8}yh+O(g^2)\ , \tag{6.9b}$$

$$b\frac{d\gamma_t}{db}=y\frac{\gamma_t}{8}+O(g^2)\ , \tag{6.9c}$$

$$b\frac{dy}{db}=-\varepsilon y+\frac{y^2}{8}+O(g^3)\ . \tag{6.9d}$$

Equations (6.9) suggest a nontrivial fixed point at $g^*=0$, $\gamma_t^*=\infty$, and $y^*=g^*\gamma_t^*=8\varepsilon+O(\varepsilon^2)$, which we shall refer to as the one-loop fixed point. Linearization of Eq. (6.9d) around this fixed point leads to a relevant RG eigenvalue that gives the inverse correlation length exponent $\nu=1/\varepsilon+O(1)$. Note that, in this limit, Eq. (6.8g) gives $\gamma_c^*=2$ and that γ_c is an irrelevant variable.

We shall not discuss further details of this fixed point because we shall see in subsequent subsections that terms at higher order in the loop expansion partially invalidate these considerations. The reason is that, whenever a physical quantity (such as γ_t) flows to infinity under RG iterations, care must be taken in drawing conclusions. Without further considerations it cannot be ruled out that the fixed-point scenario is destroyed by higher-order terms. For example, if there were a term of order $y^3\gamma_t$ in Eq. (6.9d) the one-loop fixed point would not exist. We shall see below that the fixed point suggested by Eqs. (6.9) is indeed destroyed at higher order, but only due to logarithmic terms. We shall show how to calculate the leading terms to all orders, i.e., how to control the limit $\gamma_t\to\infty$. The resulting theory will again show a fixed point that is distinct from the one-loop fixed point and that corresponds to a phase transition distinct from the metal-insulator transition. We shall also see that all higher-order corrections in Eq. (6.9a) vanish. This implies that the physical disorder is not renormalized and that we are indeed dealing with the magnetic phase transition discussed in Sec. IV.B.

b. Absence of disorder renormalizations

As already mentioned, the existence of the one-loop fixed point suggested by Eqs. (6.9) is not obvious once one considers higher-order terms. Specifically, if for $\gamma_t\to\infty$ a term of order g^n grew like γ_t^p with $p>n-1$ in Eq. (6.9a), or with $p>n$ in Eq. (6.9b), or with $p>n+1$ in Eq. (6.9c), then the one-loop fixed-point scenario would break down. In fact, it is easy to find diagrammatic contributions that *appear* not to be consistent with the fixed-point scenario. For example, the diagram in Fig. 29 naively seems to lead to a term of order $g^3\gamma_t^4$ in Eq. (6.9a) because it involves four q^3 "interacting" vertices that, according to Eqs. (3.92), (3.114), and (3.117a), can each be of $O(K^{(t)})$. Note that for dimensional reasons $K^{(t)}$ must appear in the combination $K^{(t)}/H$. We also do not distinguish between $K^{(t)}$ or H and the respective renormalized quantities, k_t or h, since ultimately the latter appear in the theory.

To find out whether such violations of the one-loop fixed-point scenario occur is a counting problem that has been considered by Kirkpatrick and Belitz (1990a). The key idea was to prove that every diagram that naively has too many factors of k_t contains, upon closer examination, enough triplet propagators (which are proportional to k_t^{-1} in the limit $k_t\to\infty$) to cancel the excess k_t factors. A number of subtleties arise in this counting process, the most obvious of which is that the limit $K^{(t)}\to\infty$ does not necessarily commute with sums over frequencies and integrals over wave numbers. The computational details can be found in the reference mentioned above. The conclusions reached were as follows:

(1) To leading order in $1/\gamma_t$ for $\gamma_t\to\infty$ it is sufficient to keep terms up to order q^4 in the q expansion of the action, Eq. (3.119). Furthermore, for a calculation of skeleton diagrams to leading order the terms of $O(q^3)$ in the action always have to be contracted in a particular way, such that two factors of S_3 can be combined into an effective term of $O(q^4)$. The effective action in the limit $\gamma_t\to\infty$ can then be written

$$S^{\rm eff}=S_2[q]+S_4^0[q]+S_4^{i,\rm eff}[q]\ . \tag{6.10}$$

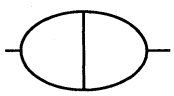

FIG. 29. Two-loop contribution to $\Gamma^{(2)}$ which, according to naive counting, is of order $(K^{(t)})^4$.

For the calculation of insertion diagrams special care has to be used, since contributions occur that are not contained in Eq. (6.10). This does not constitute a serious problem, though, since all insertions can be obtained recursively from lower-order skeleton diagrams. In Eq. (6.10) S_2 is the Gaussian action, Eqs. (3.120), and S_4^0 is the term of $O(q^4)$ that results from the noninteracting part of the action, Eq. (3.118) with $S_{\rm int}=0$. $S_4^{i,\rm eff}$ is a combination of the term of $O(q^4)$ in $S_{\rm int}$ and the mentioned partial contraction of $(S_3)^2$. It can be written as

$$S_4^{i,\rm eff}=\frac{\pi T K^{(t)}}{16}\sum_{\substack{n_1 n_2\\ n_3 n_4}}\sum_{\alpha}\sum_{i=1}^{3}\sum_{r=0,3}(-1)^r\delta_{n_1+n_3,n_2+n_4}$$
$$\times\int_{\mathbf{p}_1,\ldots,\mathbf{p}_4}(2\pi)^d\delta(\mathbf{p}_1+\mathbf{p}_2+\mathbf{p}_3+\mathbf{p}_4)\Delta(\mathbf{p}_1+\mathbf{p}_2,|\Omega_{n_2}-\Omega_{n_1}|)$$
$$\times[\mathrm{tr}[\tau_r\otimes s_i(q(\mathbf{p}_1)q^+(\mathbf{p}_2))]^{\alpha\alpha}_{n_1 n_2}-\mathrm{tr}[\tau_r\otimes s_i(q^+(\mathbf{p}_1)q(\mathbf{p}_2))]^{\alpha\alpha}_{n_1 n_2}]$$
$$\times[\mathrm{tr}[\tau_r\otimes s_i(q(\mathbf{p}_3)q^+(\mathbf{p}_4))]^{\alpha\alpha}_{n_3 n_4}-\mathrm{tr}[\tau_r\otimes s_i(q^+(\mathbf{p}_3)q(\mathbf{p}_4))]^{\alpha\alpha}_{n_3 n_4}]\,, \tag{6.11a}$$

with

$$\Delta(\mathbf{p}_1+\mathbf{p}_2,|\Omega_{n_2}-\Omega_{n_1}|)=1-\frac{K^{(t)}|\Omega_{n_2}-\Omega_{n_1}|}{(\mathbf{p}_1+\mathbf{p}_2)^2/G+(K^{(t)}+H)|\Omega_{n_2}-\Omega_{n_1}|}\,. \tag{6.11b}$$

Note that

$$\lim_{K^{(t)}\to\infty}\Delta(\mathbf{p}_1+\mathbf{p}_2,|\Omega_{n_2}-\Omega_{n_1}|)=0\,, \tag{6.12}$$

so that formally $S_4^{i,\rm eff}$ is of order $(K^{(t)})^0$. However, depending on the quantity to be calculated, this conclusion may turn out to be incorrect if the limit is taken *after* doing the integrals. Similarly, S_4^0, which is of $O(q^4)$, can produce factors of $K^{(t)}$ upon integration if dimensional regularization is used.

(2) For the renormalization of G neither the noncommutability nor any of the other problems mentioned occur. As a result, the correction terms in Eq. (6.9a) have been shown to vanish to all orders. That is, the physical disorder is not renormalized, and the exact solution of the g-flow equation to leading order for large γ_t is

$$g=b^{-\varepsilon}g(b=1)\,. \tag{6.13}$$

(3) The one-point vertex function $\Gamma^{(1)}$ is not renormalized either. This implies that in the limit $\gamma_t\to\infty$ there is no wave-function renormalization, $\mathbb{Z}=1$.

(4) For H and $K^{(t)}$, Eq. (6.12) does *not* imply the absence of renormalizations, as can be seen already at one-loop order [see Eqs. (6.9b)–(6.9d)]. Still, in the flow equations for h, γ_t, and y the conditions for the existence of the fixed point mentioned at the beginning of the present subsection are fulfilled up to possible logarithmic terms to which the counting arguments are not sensitive. Logarithmic terms that violate the fixed-point conditions do indeed occur, starting at two-loop order; see the next subsection. Kirkpatrick and Belitz (1990a) have pointed out that the *renormalized* vertex functions at one-loop order already suggest that there will be logarithmic problems with the fixed point at higher order.

c. Two-loop results

To leading order for $\gamma_t\to\infty$ the vertex functions $\Gamma_0^{(2)}$ and $\Gamma_t^{(2)}$ corresponding to the two-point propagators defined in Eqs. (5.2b) and (5.17), respectively, have been calculated to two-loop order (Kirkpatrick and Belitz, 1989, 1990a). Logarithmically divergent terms were found that invalidate the one-loop fixed point suggested by Eqs. (6.9).

The renormalization of H is determined by $\Gamma_0^{(2)}$. The calculation is greatly facilitated by use of the effective action, Eqs. (6.10) and (6.11). Equation (6.12) turns out to be misleading in this case, since after doing the frequency sums one finds a contribution to leading order in $K^{(t)}$. Also, an insertion diagram has to be considered separately. With dimensional regularization, the result of the calculation is

$$\Gamma_0^{(2)}(\mathbf{k},\Omega_m)=\frac{\mathbf{k}^2}{G}+H\Omega_m-H\Omega_m\frac{3\bar{G}}{8\varepsilon}\left[\frac{K^{(t)}}{H}\right]+H\Omega_m\frac{3}{32\varepsilon^2}\left[\frac{K^{(t)}}{H}\right]^2\bar{G}^2[1+\varepsilon L_t+\varepsilon\ln(GH\Omega_m)-3\varepsilon/2]+O(G^2\varepsilon^0,G^3)\,. \tag{6.14}$$

The $K^{(t)}$ renormalization can be computed by considering $\Gamma_t^{(2)}$. Again, the effective action can be utilized if certain precautions are taken, and the procedure is explained in detail in the original literature. To leading order for large $K^{(t)}$, and in dimensional regularization, the result is

$$\Gamma_t^{(2)}(\mathbf{k},\Omega_m)=\frac{\mathbf{k}^2}{G}+(H+K^{(t)})\Omega_m-K^{(t)}\Omega_m\frac{G}{2\varepsilon}\left[\frac{K^{(t)}}{H}\right] +K^{(t)}\Omega_m\frac{5\bar{G}^2}{32\varepsilon^2}\left[\frac{K^{(t)}}{H}\right]^2\left\{1+\varepsilon\left[-\frac{3}{10}L_t-\frac{1}{2}L_s\ln 2+\ln(GH\Omega_m) -\frac{1}{4}+\frac{\pi^2}{48}+\frac{3}{10}\ln 2+\frac{1}{4}(\ln 2)^2\right]\right\}+O(\varepsilon^0,\bar{G}^3)\ . \tag{6.15}$$

Equations (6.14) and (6.15) explicitly confirm the absence of a disorder renormalization at two-loop order. They also show the logarithmic terms that will violate the fixed-point scenario discussed earlier. We note, however, that not all logarithms that show in the perturbative expressions appear in the flow equations. For instance, the contribution proportional to L_s in the two-loop term in Eq. (6.15) is canceled in deriving the flow equations.

Now we can use the renormalization procedure described in Sec. VI.A.1.a above to determine the RG flow equations to two-loop order. From Sec. VI.A.1.b, result (3), we know that there is no wave-function renormalization. From Eqs. (6.7) together with Eqs. (6.14) and (6.15) the renormalization constants to leading order for $K^{(t)}\to\infty$ are

$$\mathbb{Z}_h=1+\frac{3}{8\varepsilon}y+\frac{3}{32\varepsilon^2}y^2(1+\varepsilon)+O(g^3)\ , \tag{6.16a}$$

$$\mathbb{Z}_t=1+\frac{1}{2\varepsilon}y+\frac{5}{32\varepsilon^2}y^2[1+\varepsilon(\tfrac{3}{10}l_t-\tfrac{3}{20})]+O(g^2)\ . \tag{6.16b}$$

This yields the following flow equations (Kirkpatrick and Belitz, 1990a):

$$b\frac{dh}{db}=h\frac{3}{8}y+h\frac{3}{16}y^2+O(g^3)\ , \tag{6.17a}$$

$$b\frac{d\gamma_t}{db}=\gamma_t\frac{1}{8}y+\gamma_t\frac{3}{32}y^2l_t-\gamma_t\frac{15}{64}y^2+O(g^3)\ , \tag{6.17b}$$

$$b\frac{dy}{db}=-\varepsilon y+\frac{1}{8}y^2+\frac{3}{32}y^3l_t-\frac{15}{64}y^3+O(g^4)\ . \tag{6.17c}$$

Comparing with the criteria given in Sec. VI.A.1.b, we see that Eq. (6.17a) is consistent with the fixed-point scenario, but Eqs. (6.17b) and (6.17c) are not. The reason is the logarithmic term at two-loop order. The mathematical origin of this term is that the limit $K^{(t)}\to\infty$ does not commute with performing frequency sums. We note that the counting arguments used in Sec. VI.A.1.b are insensitive to these terms, so that they can only be found by an explicit calculation.

d. Resummation of leading singularities

In the previous subsection we showed that the one-loop fixed point proposed in Sec. VI.A.1.a does not exist, due to logarithmically divergent terms in the RG flow equations. An inspection of the general structure of the diagrams shows that one should expect increasing powers of logarithms to appear at higher orders in the loop expansion. The leading term of n-loop order in Eq. (6.17c) is expected to be proportional to $y^{n+1}l_t^{n-1}$ unless cancellations occur. This raises the question of whether one can control these divergencies by performing a resummation to all orders in the loop expansion. This question has been answered affirmatively (Kirkpatrick and Belitz, 1990b; Belitz and Kirkpatrick, 1991). In this subsection we describe the resummation. In Sec. VI.A.2 we shall see that the resulting theory contains the pseudomagnetic phase transition discussed in Sec. IV.B.

To motivate the resummation discussed above we first explicitly consider the wave-number- and frequency-dependent one-loop corrections to G, H, and $K^{(t)}$,

$$\delta G^{(1\text{-loop})}(\mathbf{k},\Omega_n)=O(G(K^{(t)})^0)\ , \tag{6.18a}$$

$$\delta H^{(1\text{-loop})}(\mathbf{k},\Omega_n)=\delta H^{(1\text{-loop})}(\Omega_n)=\frac{3}{4}GK^{(t)}\frac{2\pi T}{\Omega_n}\sum_{l=1}^{n}\left[1-\frac{\Omega_l}{\Omega_n}\right]\int_{\mathrm{p}}\mathcal{D}_l^t(\mathbf{p})+O(G(K^{(t)})^0)\ , \tag{6.18b}$$

$$\delta K^{(t),(1\text{-loop})}(\mathbf{k},\Omega_n)=G^2(K^{(t)})^2\pi T\sum_{l=1}^{\infty}\int_{\mathrm{p}}\mathcal{D}_{l+n}(\mathbf{p}+\mathbf{k})\mathcal{D}_l(\mathbf{p})+O(G(K^{(t)})^0)\ . \tag{6.18c}$$

Here we have neglected terms that vanish in the limit $g\to 0$, $\gamma_t\to\infty$, $y=g\gamma_t$ finite. We have also neglected a term in Eq. (6.18c) which is UV finite and was found in the original literature to be of no qualitative importance. Performing the integrals in Eqs. (6.18) one recovers the one-loop terms in Eqs. (6.14) and (6.15).

In order to obtain the generalization of Eq. (6.18c) to all orders, we first reconsider the effective action, Eq.

(6.10). The term $S_4^{i,\mathrm{eff}}$ consists of the original four-point vertex in S_{int} and two of the original three-point vertices, which have been partially contracted by means of a triplet propagator. A diagrammatic representation is shown in Fig. 30(a). The two terms on the right-hand side of this diagrammatic equality correspond to the two terms on the right-hand side of Eq. (6.11b). The noninteracting four-point vertex S_4^0 does not contribute to the $K^{(t)}$ renormalization in the limit considered. In terms of these diagrams the one-loop contribution to $\Gamma_t^{(2)}$, Eq. (6.18c), is simply a bubble; see Fig. 30(b). At two-loop order, the two relevant skeleton diagrams in the limit $K^{(t)}\to\infty$ are shown in Fig. 30(c). The diagram shown in Fig. 30(d) contains both of these as well as a reducible diagram that does not contribute to $\Gamma_t^{(2)}$. Since according to Eq. (6.12) the effective vertex is of $O((K^{(t)})^0)$ this means that, to leading order for $K^{(t)}\to\infty$, the two irreducible skeleton diagrams in Fig. 30(c) can be replaced by minus the reducible diagram in Fig. 30(d) [note that, in this case, no integrations over the arguments of $\Delta(\mathbf{p},\Omega)$, Eq. (6.11b), are involved]. More generally, at n-loop order this equality can be generalized so that the sum of all relevant irreducible diagrams can be related to reducible diagrams. Diagrammatically this is shown in Fig. 30(e). In the limit $K^{(t)}\to\infty$ the dashed line on the right-hand side of the equality should be replaced by unity; cf. Eq. (6.11b). Technically, the geometric series shown in Fig. 30(e) can be summed by replacing one of the $K^{(t)}$ factors in Eq. (6.18c) by $K^{(t)}(\mathbf{k},\Omega_n)$. This sums all skeleton diagrams. Insertions can be trivially included by dressing the two propagators in the one-loop bubble. Analytically continuing the resulting equation to real frequencies and taking the zero-temperature limit yields

$$D_s(\mathbf{k},\Omega)=D_s^0+\frac{iG}{2}\int_{\mathrm{p}}\frac{1}{\mathbf{p}^2(\mathbf{p}+\mathbf{k})^2}\times\int_0^\infty d\omega\,\chi(\mathbf{p},\omega)\chi(\mathbf{p}+\mathbf{k},\omega+\Omega)\ , \tag{6.19a}$$

where $D_s(\mathbf{k},\Omega)$ is the spin diffusion coefficient [cf. Eq. (3.128b)],

$$D_s(\mathbf{k},\Omega)=1/GK^{(t)}(\mathbf{k},\Omega)\ , \tag{6.19b}$$

with D_s^0 the bare D_s. χ in Eq. (6.19a) is the retarded susceptibility,

$$\chi(\mathbf{p},\Omega)=\mathbf{p}^2/[-i\Omega/D(\Omega)+\mathbf{p}^2]\ . \tag{6.20a}$$

$D(\Omega)$ is the dressed, frequency-dependent heat diffusivity, Eq. (3.128c),

$$D(\Omega)=1/GH(\Omega)\ , \tag{6.20b}$$

where we have anticipated that D will not depend on wave number.

An equation for $D(\Omega)$ can be obtained by dressing the lowest-order perturbative result for H, Eq. (6.18b). Accordingly, we dress the triplet propagator in Eq. (6.18b). We also must dress the factor of $K^{(t)}$. It is important to note that this $K^{(t)}$ is not a vertex, but originates from a propagator via the identity

$$\mathbf{p}^2\mathcal{D}_n^t(\mathbf{p})=1-K^{(t)}G\Omega_n D_n^t(\mathbf{p})\ , \tag{6.21}$$

as can be verified from the perturbation theory. It therefore has to be dressed under the integral. One finds

$$\frac{1}{D(\Omega)}=\frac{1}{D^0}+\frac{3}{4}G\int_{\mathrm{p}}\frac{1}{\Omega}\int_0^\Omega d\omega(1-\omega/\Omega)\times\frac{1}{-i\omega+\mathbf{p}^2D_s(\mathbf{p},\omega)}\ , \tag{6.22}$$

with $D^0=1/GH$.

Equations (6.19), (6.20), and (6.22) form a closed set of integral equations for the diffusion coefficients D_s and D. Note that Eqs. (6.19) for D_s are exact in the limit $K^{(t)}\to\infty$ or $D_s\to 0$, since only the diagrams shown in Fig. 30(e) contribute in this limit. Equation (6.22) for D correctly reproduces perturbation theory up to two-loop order, as one can see by iterating Eq. (6.22) and comparing with Eq. (6.14). The quality of the expression for D, Eq. (6.22), beyond two-loop order is not known. We also note that both equations have been derived only in the

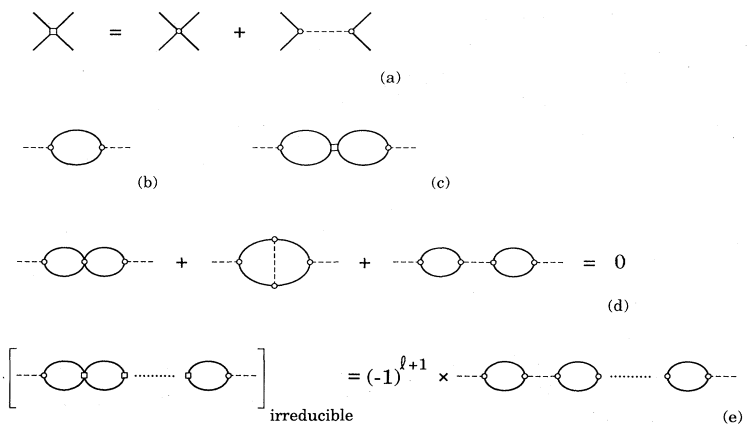

FIG. 30. Resummation of the perturbation expansion for the triplet vertex function. (a) Effective four-point vertex (square) in terms of original three- and four-point vertices (circles). Dashed lines denote triplet propagators. (b) Leading one-loop contribution to $\Gamma_t^{(2)}$. (c) Leading two-loop skeleton contributions to $\Gamma_t^{(2)}$. (d) Cancellation of reducible and irreducible diagrams at two-loop order. (e) Relation between reducible and irreducible diagrams at n-loop order.

limit of small $|\mathbf{k}|$ and Ω. Outside of this regime there may be arbitrary corrections over which we have no control. These corrections are irrelevant for universal quantities such as critical exponents, but they are important for nonuniversal quantities. We shall come back to this point later.

2. The pseudomagnetic phase transition

The integral equations (6.19) and (6.22) have been solved numerically and by construction of an analytic scaling solution (Kirkpatrick and Belitz, 1990b; Belitz and Kirkpatrick, 1991). This work also contained a tentative RG analysis, which later was improved upon (Kirkpatrick and Belitz, 1992b). All three approaches show a phase transition where the spin diffusivity $D_s(0,0)$ vanishes, while the specific heat shows a logarithmic nonanalyticity, and the charge diffusivity is uncritical. The two analytic approaches further show that the critical behavior is unconventional, i.e., nonpower law, in the (unobservably small) critical region. The numerical and the two analytical approaches represent solutions of increasing technical complexity, and we shall discuss them in this order.

a. Numerical solution

The first point to notice for a solution of Eqs. (6.19) and (6.22) is that the integrals in these equations do not exist unless there is an ultraviolet cutoff on the momentum integrals. The universal critical behavior to be obtained below will of course not depend on this cutoff. For the sake of mathematical convenience we implement a simple soft cutoff by multiplying the wave-number integrands in Eqs. (6.19) and (6.22) by

$$f(\mathbf{p})=q_0^2/(q_0^2+\mathbf{p}^2)\ . \tag{6.23}$$

For reasons explained in Sec. II.A.1.d we expect the momentum cutoff q_0 to be on the order of the Fermi momentum k_F. The fact that in the original theory one had $f(\mathbf{p})=1$ reflects again the fact that the derivation of the equations is valid only for small wave numbers. With Eq. (6.23) the wave-number integral in Eq. (6.19a) can be performed analytically by means of a Feynman parametrization trick (Feynman, 1949).

The remaining integral equations can be solved by iteration. It is convenient to introduce a disorder-related coupling constant $\lambda=Gk_F/4\pi$ and to measure diffusivities in units of $1/2m$. In these units, the bare diffusivities D^0 and D_s^0 can be expressed in terms of Fermi-liquid parameters, (see Sec. III.B.3.d),

$$D^0=\frac{2/\lambda}{1+F_1^s/3}\ , \tag{6.24a}$$

$$D_s^0=D^0(1+F_0^a)\ . \tag{6.24b}$$

The integral equations can now be solved for given values of λ, F_0^s, and F_0^a. Figure 31 shows the result of such an iterative solution[17] for $d=3$, $F_0^s=0$, $F_1^s=0$, and $F_0^a=-0.9$. We note the following features of the numerical solution: (1) For the parameters chosen, D_s vanishes at a finite critical value of the disorder parameter, $\lambda=\lambda_c$. This supports the idea of a sharp phase transition that is distinct from the metal-insulator transition; (2) if $t=(\lambda_c-\lambda)/\lambda_c$ is the dimensionless distance from this critical point, then $D_s\equiv D_s(0,0)\sim t^{s_s}$ with $s_s=\nu(\kappa+\varepsilon)$ [see Eq. (4.30a)], very close to one; (3) $D\equiv D(0)$ has a much weaker singularity than D_s. The behavior of D is consistent with a power-law $D\sim t^{s_h}$ with an apparent exponent $s_h=\nu\kappa\simeq 0.4$; (4) for $|F_0^a|<|(F_0^a)_c|$, D_s was not found to go to zero and the iterative method failed to converge at large values of λ. The value of $(F_0^a)_c$ was found to depend strongly on the value of the momentum cutoff q_0 and also on the shape of the cutoff function. For the cutoff given by Eq. (6.23) with $q_0=2k_F$, $(F_0^a)_c\cong 0.7$. We shall come back to points (3) and (4) after an analytic discussion of the equations.

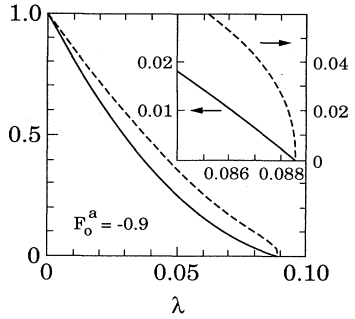

FIG. 31. Numerical solution of Eqs. (6.19) and (6.22). Shown are the normalized spin diffusivity D_s/D_s^0 (solid line), and heat diffusivity D/D^0 (dashed line) vs the disorder parameter λ. Fermi-liquid parameters have been chosen as $F_0^s=0$, $F_0^a=-0.9$. From Belitz and Kirkpatrick (1991).

b. Analytic scaling solution

The numerical solution of Eqs. (6.19) and (6.22) discussed in the previous subsection describes a sharp phase transition in at least some parts of parameter space. This motivates constructing a scaling or similarity solution to these equations. Here we do so, restricting ourselves to $2<d<4$.

In Eqs. (6.19) and (6.22) let the critical disorder be G_c and the distance from criticality $t=|G_c-G|/G_c$. At

[17]The numerical result shown has been obtained by solving equations that are slightly more complicated than Eqs. (6.19) and (6.22), since they contain the additional contribution mentioned in connection with Eqs. (6.18). This leads to only minor quantitative differences.

$t=0$ insert the ansatz [see Sec. IV.B, Eqs. (4.27b)–(4.27d)]

$$D(\Omega\to 0)\simeq a\Omega^{\kappa/z},\quad t=0\ , \tag{6.25a}$$

$$D_s(\mathbf{k},\Omega)\simeq a_s\Omega^{(\varepsilon+\kappa)/z}+b_s k^{(\kappa+\varepsilon)},\quad t=0\ , \tag{6.25b}$$

into Eqs. (6.19) and (6.22). Using asymptotic analysis, one finds

$$\kappa=0\ , \tag{6.26a}$$

$$z=2\ , \tag{6.26b}$$

except for logarithmic corrections. At the same level of approximation the structure of Eq. (6.19) gives

$$D_s(t,\mathbf{k}=0,\Omega=0)\sim t^{\nu\epsilon}\ , \tag{6.27a}$$

with

$$\nu=\frac{1}{d-2}\ . \tag{6.27b}$$

To characterize the logarithmic corrections to this scaling behavior, we generalize the above ansatz to

$$D(\Omega)=D^0/F(\ln(1/\Omega\tau)),\quad t=0\ , \tag{6.28a}$$

$$D_s(\mathbf{k},\Omega=0)=D_s^0|\mathbf{k}/k_F|^{d-2}F_s(-\ln|\mathbf{k}/k_F|),\quad t=0\ , \tag{6.28b}$$

and similarly for $t\neq 0$ and $\Omega\neq 0$ (for D_s). In Eqs. (6.28) τ is a scattering rate and k_F is the Fermi wave number. F and F_s can be determined by using Eqs. (6.28) in Eqs. (6.19) and (6.22). After some algebra, a soluble functional differential equation is obtained for these quantities. Asymptotically close to the critical point one obtains (Kirkpatrick and Belitz, 1992b)

$$D(\Omega\to 0)=D^0\left[\frac{2}{\pi}\ln(d/2)\right]^{1/2}\exp\left\{-\left[\ln\left[c(d)\ln\frac{1}{\Omega\tau}\right]\right]^2\Big/2\ln(d/2)\right\}\ , \tag{6.29}$$

$$D_s(\mathbf{k}\to 0,\Omega=0)=D_s^0|\mathbf{k}/k_F|^{d-2}\tilde{d}Gk_F^{d-2}\gamma_t^0\left[\frac{2}{\pi}\ln(d/2)\right]^{1/2}\exp\{-[\ln[c(d)\ln|\mathbf{k}/k_F|^{-1}]]^2/2\ln(d/2)\}\ , \tag{6.30}$$

with $c(d)$ and $\tilde{d}$ nonsingular functions of d. We note that the critical behavior of D is in between a power-law and logarithmic behavior in the sense that $D(\Omega\to 0)$ vanishes more slowly than any power of Ω, yet faster than any power of $1/\ln\Omega$. Equations (6.29) and (6.30) are both valid at the critical point, $t=0$. Using similar techniques, one can determine D and D_s as functions of t, or $D_s(\mathbf{k}=0,\Omega)$ at $t=0$. General scaling arguments (see Sec. IV) suggest that the structure of the result would be the same as in Eqs. (6.29) and (6.30), with, for example, t replacing $\Omega\tau$ in Eq. (6.29). This is indeed the case, as will become clear from the RG analysis in the next subsection.

c. Renormalization-group analysis

The question arises of how the asymptotic solution provided in the previous section can be related to the various RG analyses in Secs. VI.A.1.a and VI.A.1.c. For this purpose it is useful to return to the perturbation theory, of which the integral equations given by Eqs. (6.19) and (6.22) constitute a resummation. We also must remember that the RG applies to the coupling constants of the underlying field theory rather than to the diffusivities of Sec. VI.A.1.d. We recall that the former are related to the latter by

$$H=1/GD^0\ , \tag{6.31}$$

$$K_t=-H+1/GD_s^0\ . \tag{6.32}$$

Between the renormalized quantities there are analogous relations (see Sec. III.B.3.d).

In the field theory, $H=\pi N_F/2$ is the bare frequency or temperature renormalization factor, and $K^{(t)}$ is the bare interaction parameter for the spin-triplet channel. The loop expansion is a perturbation theory in powers of $Y=GK^{(t)}/H$. We return from the real-frequency formulation of the previous subsections to the original Matsubara formulation with Matsubara frequencies $\Omega_n=2\pi Tn$. The dressed, frequency- and wave-number-dependent counterparts of H and K we denote by $h(n)$ and $k_t(\mathbf{p},n)$, respectively. Equations (6.19) and (6.22) are then equivalent to the following equations for h and k_t:

$$\frac{h(n)}{H}=1+\frac{3}{2}Y\int_{\mathrm{p}}\frac{1}{n}\sum_{l=1}^{n}(1-l/n)\frac{k_t(\mathbf{p},l)}{K^{(t)}}\frac{1}{\mathbf{p}^2\Omega_l[h(l)+k_t(\mathbf{p},l)]G}\ , \tag{6.33}$$

$$\frac{k_t(\mathbf{q},n)}{K^{(t)}}=1+YH\frac{k_t(\mathbf{q},n)}{K^{(t)}}\int_{\mathrm{p}}2\pi T\sum_{l=1}^{\infty}\frac{1}{\mathbf{p}^2+\Omega_l h(l)G}\frac{1}{(\mathbf{p}+\mathbf{q})^2+(\Omega_l+\Omega_n)h(l+n)G}\ . \tag{6.34}$$

Instead of the cutoff regularization of the previous subsections we shall employ dimensional regularization here, and have therefore omitted the cutoff function $f(\mathbf{p})$.

Iteration of Eqs. (6.33) and (6.34) generates the loop expansion, an expansion of h and k_t in powers of Y. It is instructive first to consider the low-order terms in this expansion. To second order in Y we find

$$h(n)/H=1-Y\frac{3}{8\varepsilon}[1+O(\varepsilon)] + Y^2\frac{3}{32\varepsilon^2}(GH\Omega_n)^{\varepsilon}(1+\gamma_t^0)^{\varepsilon} \times[1-3\varepsilon/2+O(\varepsilon^2)] , \tag{6.35}$$

$$k_t(\mathbf{q}=0,n)/K^{(t)}=1-\frac{Y}{2\varepsilon}[1+O(\varepsilon)] + Y^2\frac{5}{32\varepsilon^2}(GH\Omega_n)^{\varepsilon} \times[1-\varepsilon(\tfrac{1}{4}+\tfrac{3}{10}L_t)+O(\varepsilon^2)] , \tag{6.36}$$

where $\gamma_t^0=K^{(t)}/H$. Equations (6.35) and (6.36) have the same structure as the two-loop perturbative result derived in Sec. VI.A.1.c. The discrepancy at two-loop order between Eq. (6.36) and Eq. (6.15) is due to the term omitted from Eqs. (6.18c) and (6.19a). Renormalization by means of minimal subtraction leads to a flow equation, Eq. (6.17c), for y, the scale-dependent counterpart of Y. To one-loop order Eq. (6.17c) seems to allow for a fixed point $y^*=4\varepsilon$, but this prospect is ruined by the $y^3 l_t$ term at two-loop order. As mentioned in Sec. VI.A.1.d, these logarithmic terms also appear at higher-loop order with the leading term at n-loop order of order $y^{n+1}l_t^{n-1}$. Yet the full equations, which resum the loop expansion to all orders, allow for a phase transition where $y=$const. It is then natural to assume that a resummation of *all* logarithmic terms would result in a function of l_t that tends to a finite value as $l_t\to\infty$.

The above scenario has been verified by a detailed calculation (Kirkpatrick and Belitz, 1992b). For $l_t\to\infty$ the results can be summarized by the RG flow equations,

$$b\frac{dy}{db}=-\varepsilon y+\frac{y^2}{2\alpha\alpha'[1-(1-1/\alpha')y/2\alpha\varepsilon]} , \tag{6.37}$$

$$b\frac{dh}{db}=\frac{(\alpha-1)yh}{2\alpha\alpha'[1-(1-1/\alpha')y/2\alpha\varepsilon]} . \tag{6.38}$$

In these equations α and α' characterize the solution of differential equations that resum perturbation theory. They satisfy the transcendental equations,

$$\alpha=\tfrac{4}{3}(\alpha-1)\tilde{\gamma}_t^{(\alpha-1)\varepsilon/2} , \tag{6.39a}$$

$$\alpha'=\alpha-\frac{\alpha(\alpha-1)\tilde{\gamma}_t^{-\alpha'\varepsilon/2}}{\alpha-1+\alpha'} , \tag{6.39b}$$

with

$$\tilde{\gamma}_t=(1+\gamma_t)e^{-3/2} . \tag{6.39c}$$

The factor $e^{-3/2}$ in Eq. (6.39c) can actually be replaced by unity if one wants to keep leading logarithmic terms only. It is retained because it contributes to the finite term $-(15/64)y^3$ in Eq. (6.17c). Terms like this turn out to have a qualitative effect on the phase diagram even though they are not important for the Fermi-liquid–incompletely frozen spin (FL-IFS) phase transition itself. For this qualitative effect the exact value of the exponent ($-\frac{3}{2}$ above) is insignificant; it is only important that it is negative.

Note that, for $\varepsilon=0$, Eqs. (6.39) give $\alpha=4$ and $\alpha'=1$. Use of these values in Eqs. (6.37) and (6.38) leads to the one-loop RG flow equations discussed in Sec. VI.A.1.a. We further note that the $\varepsilon\to 0$ approximation breaks down when $\varepsilon\ln\tilde{\gamma}_t\cong\varepsilon l_t\gtrsim 1$, as expected from the discussion below Eq. (6.36).

Before proceeding let us briefly discuss in more detail the transcendental equations given by Eqs. (6.39a) and (6.39b). First note that $\tilde{\gamma}_t$ is given in terms of $\gamma_t=y/g=y/g(b=1)b^{-\varepsilon}$ [cf. Eq. (6.13)]. At $b=1$, $\tilde{\gamma}_t(b=1)=\tilde{\gamma}_t^0$, and for future use it is important to note that Eqs. (6.39) should also be satisfied for b close to 1. Here we discuss these equations for general b. Equation (6.39a) always has one and only one solution if $\tilde{\gamma}_t\geq 1$. $\alpha=4$ for $\tilde{\gamma}_t=1$ and $\alpha\to 1$ for $\tilde{\gamma}_t\to\infty$. If $(\tilde{\gamma}_t)^c<\tilde{\gamma}_t<1$ with $(\tilde{\gamma}_t)^c=\exp[-1/\varepsilon\times 0.0446\ldots]$ then there are two solutions, of which the one yielding the larger α is by continuity the physical one. For $\tilde{\gamma}_t=(\tilde{\gamma}_t)^c$ there is one solution, giving the largest possible $\alpha=\alpha_c=7.218\ldots$. For $\tilde{\gamma}_t<(\tilde{\gamma}_t)^c$ there is no solution. The physical meaning of this absence of a solution is that, if the bare value γ_t^0 of γ_t is smaller than a critical value $(\gamma_t^0)^c$, then the assumption of γ_t scaling to infinity fails, and there is no FL-IFS transition. We shall come back to this in the next subsection. α' vanishes for $\tilde{\gamma}_t=(\tilde{\gamma}_t)^c$, then increases with increasing $\tilde{\gamma}_t$, goes through a maximum, and approaches α for large $\tilde{\gamma}_t$. Both α and α' as functions of $\tilde{\gamma}_t$ show square-root cusps at $(\tilde{\gamma}_t)^c$. Note that these considerations imply that at the expected fixed point $\alpha=\alpha'=1$. Naively using this result in Eqs. (6.37) and (6.38) leads to the exponents given by Eqs. (6.26a) and (6.27b). To recover the logarithmic corrections discussed in the previous subsection, one must treat the approach of α to its asymptotic value carefully.

The analysis proceeds by first defining the RG β function. Equation (6.37) yields

$$\beta(y)=b\frac{dy}{db}=-\varepsilon y\frac{1-y/y^*}{1-(1-1/\alpha')y/y^*} . \tag{6.40}$$

$\beta(y)$ has a zero at $y=y^*\equiv 2\alpha\varepsilon$, as it should, and its derivative determines the correlation length exponent ν,

$$\nu=1/\beta'(y^*)=1/\alpha'\varepsilon . \tag{6.41}$$

Asymptotically, we have again $\nu=1/\varepsilon$. For h and k_t, we find for $y=y^*$ from Eqs. (6.37) and (6.38) the flow equations

$$b\frac{d\ln h}{db}=\varepsilon(\alpha-1)\ , \tag{6.42}$$

$$b\frac{d\ln k_t}{db}=\varepsilon\alpha\ . \tag{6.43}$$

For γ_t this yields

$$b\frac{d\ln\gamma_t}{db}=\varepsilon\ . \tag{6.44}$$

Equation (6.39a) shows that, for $\tilde{\gamma}_t\to\infty$, α approaches its asymptotic value logarithmically,

$$\alpha=1+\frac{2}{\varepsilon\ln\tilde{\gamma}_t}\ln\ln\tilde{\gamma}_t\ . \tag{6.45}$$

To leading order we have $\ln\tilde{\gamma}_t\simeq\ln\gamma_t$. Combining Eqs. (6.42), (6.44), and (6.45), we find for $b\to\infty$

$$b\frac{d\ln h}{db}=\frac{2}{\varepsilon\ln b}\ln(\ln b)\ , \tag{6.46}$$

with the solution

$$h(b)\sim\exp\left[\frac{1}{\varepsilon}[\ln(\ln b)]^2\right]\ . \tag{6.47}$$

The same method yields for k_t

$$k_t(b)\sim b^{\varepsilon}\exp\left[\frac{1}{\varepsilon}[(\ln(\ln b)]^2\right]\ . \tag{6.48}$$

We see that, in the case of h, the asymptotic exponent $\alpha-1=0$ actually implies the log-log normal behavior expressed by Eq. (6.47), and in the case of k_t, there is a corresponding correction to the naive power law. The scale-dependent diffusion coefficients are given by Eqs. (6.47) and (6.48) via $D(b)\sim 1/h(b)$ and $D_s(b)\sim 1/k_t(b)$. Finally, the dependence on the physical parameters t, frequency Ω, and wave number $|\mathbf{p}|$ in the scaling regime can be obtained from the relations

$$\ln b=\nu\ln(1/t)\ , \tag{6.49}$$

for the behavior of the static, homogeneous diffusivities as a function of t, and

$$\ln b=\frac{1}{z}\ln(1/\Omega\tau)\ , \tag{6.50}$$

$$\ln b=-\ln|\mathbf{p}/k_F|\ , \tag{6.51}$$

for the momentum and frequency dependence at the critical point. Asymptotically, we have $\nu=1/\varepsilon$, $z=2$, and $\kappa=0$. Comparison with Sec. VI.A.2.b shows that the RG solution is in full agreement with the asymptotic behavior of the integral equations if the latter are evaluated to leading order in ε only.

We conclude this subsection by briefly discussing the two-dimensional case. First we note that the RG calculations given so far cannot be directly applied because in the theory the limits $\varepsilon\to 0$ and $\gamma_t\to\infty$ do not commute. Further, in two-dimensional systems there is no FL-IFS phase transition, and as a consequence of this there is no physically accessible fixed point. Rather, the theory leads to a runaway RG trajectory for γ_t and y. However, by examining the perturbation theory in detail, we can make a limited amount of progress. It has been argued (Kirkpatrick and Belitz, 1992b) that, for small scales or for $y\ln(1+\gamma_t)<1$, the relevant RG flow equation is approximately

$$b\frac{dy}{db}=\frac{y^2}{4}\ . \tag{6.52}$$

For larger b, where $y\ln(1+\gamma_t)>1$, the flow equation for y is

$$b\frac{dy}{db}=y^2\ . \tag{6.53a}$$

Note that the solution of Eq. (6.53a) is

$$y(b)=\frac{y_0}{1-y_0\ln b}\ , \tag{6.53b}$$

with $y_0=y(b=1)$, which suggests that y diverges at a finite scale. In reality Eq. (6.53b) breaks down before this scale is reached, and for larger scales a theory for the IFS phase is needed. As we already mentioned, such a theory is not presently available.

d. Summary

In summary, in some parts of parameter space there is a FL-IFS phase transition, which is caused by disorder but is distinct from the metal-insulator transition. The transition is characterized by a divergent magnetic susceptibility χ, a weakly singular specific-heat coefficient γ, and noncritical charge transport. The singular functions near the phase transition satisfy the scaling laws

$$\chi(t,T)=b^{\varepsilon}\exp\left[\frac{(\ln\ln b)^2}{2\ln(d/2)}\right]\chi(b^{\varepsilon}t,b^2T)\ , \tag{6.54a}$$

$$\gamma(t,T)=\exp\left[\frac{(\ln\ln b)^2}{2\ln(d/2)}\right]\gamma(b^{\varepsilon}t,b^2T)\ , \tag{6.54b}$$

$$D_s(t,\mathbf{k},\omega)=b^{-\varepsilon}\exp\left[-\frac{(\ln\ln b)^2}{2\ln(d/2)}\right]D_s(b^{\varepsilon}t,b\mathbf{k},b^2\omega)\ , \tag{6.54c}$$

$$D(t,\mathbf{k},\Omega)=\exp\left[-\frac{[\ln(\ln b)]^2}{2\ln(d/2)}\right]D(b^{\varepsilon}t,b\mathbf{k},b^2\Omega)\ . \tag{6.54d}$$

In other parts of parameter space the situation appears different. The numerical solution method discussed in Sec. VI.A.2.a shows that in $d=3$ and for $|F_0^a|\lesssim 0.7$ there is no FL-IFS phase transition. This is consistent with the RG analysis of Sec. VI.A.2.c. In particular, in the previous subsection we noted that the transcendental equations for α and α' do not have a solution if γ_t^0 is smaller than a critical value $(\gamma_t^0)^c$. Since all of these equations

were derived under the assumption that γ_t scales to infinity (otherwise there would not be a fixed point with a nonzero y^*), an obvious interpretation of the absence of a solution is the failure of the assumption. (Other interpretations are conceivable, however; see Sec. X.B.2.a.) We conclude that the phase diagram in three dimensions looks qualitatively as shown in Fig. 28(a). For a sufficiently large value of γ_t^0, or of $|F_0^a|=\gamma_t^0/(1+\gamma_t^0)$, the system undergoes, with increasing disorder, the FL-IFS phase transition discussed in this section. With further increasing disorder, the IFS phase presumably becomes unstable, and a transition to a charge insulator occurs (we remind the reader that the IFS phase is a charge conductor). There is at present no theory for the latter transition, and the corresponding phase boundary is indicated in Fig. 28 as a dashed line. In particular, it is not known whether the dashed line will reach the line $|F_0^a|=1$ at a finite disorder. For γ_t^0 smaller than a critical value $(\gamma_t^0)^c$, γ_t does not flow to infinity with increasing scale. There consequently is no IFS phase, and it is natural to expect a metal-insulator transition instead. This is discussed in the next subsection. Of course this Fermi-liquid–insulator transition cannot be described with the theory set up here, since at the metal-insulator transition the disorder is renormalized, while here we have explicitly assumed otherwise. Furthermore, with $\gamma_t(b\to\infty)$ approaching a finite value, we lose one of our small parameters. We therefore cannot hope to find a well controlled description for arbitrary dimensionalities, as has been possible in the case of the FL-IFS phase transition. Worse than that, not even a controlled ε expansion is feasible for the Fermi-liquid–insulator transition. To see this, we consider the dependence of the critical value $(\gamma_t^0)^c$ on the dimensionality. We recall, from the discussion of Eqs. (6.39),

$$(\gamma_t^0)^c=-1+e^{3/2}\exp\left\{-\frac{1}{\varepsilon}0.0446\ldots\right\}. \tag{6.55}$$

As noted in Sec. VI.A.2.c, the factor $e^{3/2}$ in Eq. (6.55) is not exact, but rather results from a two-loop approximation for the less leading terms in the perturbation theory. The numerical value for $(\gamma_t^0)^c$ is therefore an approximation, but its ε dependence is accurate. For $d=3$, Eq. (6.55) corresponds to a critical value of F_0^a, which is $(F_0^a)_c=-0.777\ldots$. This should be contrasted with the integral equation approach, which numerically gives $(F_0^a)_c\cong-0.7$ for a particular type of momentum cutoff. However, this value is strongly cutoff dependent, and for different cutoff procedures values of $(F_0^a)_c$ as small as -0.98 have been found. It is therefore possible that the region of parameter space occupied by the IFS phase is, for 3-d systems, very small. With decreasing $\varepsilon=d-2$, this region increases, and from Eq. (6.55) we find that $(\gamma_t^0)^c$ vanishes at a value $\varepsilon=\varepsilon_0>0$ given by

$$\varepsilon_c\cong 0.030\ldots. \tag{6.56}$$

It follows that, with decreasing dimensionality, the multicritical point in the phase diagram, where the three phases meet, is shifted to smaller values of $|F_0^a|$ and reaches $F_0^a=0$ at $d_c=2+\varepsilon_c>2$. This is shown qualitatively in Fig. 28. For $d<d_c$ there is no multicritical point, and the Fermi-liquid and insulator phases are *always* separated by an IFS phase, i.e., there is no direct FL-I transition for dimensionalities close to two. An immediate consequence for the description of the FL-I transition is that one cannot treat it by means of an $\varepsilon=d-2$ expansion. Unless a different small parameter can be found, the theory of this transition is therefore necessarily nonperturbative. In the next subsection we show how one can use an approximate RG based on two-loop perturbation theory in $d=3$ and obtain approximate exponents for various transport and thermodynamic qualities. We also mention that at the multicritical point one expects a special critical behavior with crossovers to the FL-IFS critical behavior discussed above and the FL-I critical behavior discussed below, respectively.

The size of the critical region where the asymptotic behavior, Eqs. (6.29) and (6.30), or Eqs. (6.47) and (6.48), applies has also been discussed by means of the RG (Kirkpatrick and Belitz, 1992b). As a direct consequence of the logarithmically slow approach of the exponents to their asymptotic values [see Eq. (6.45)], the asymptotic critical region is exponentially small for $\varepsilon\ll 1$. To leading logarithmic accuracy one finds

$$t\le\gamma_t^0\exp\left[-\frac{3}{2}-\frac{20}{\varepsilon}\ln\frac{20}{\varepsilon}\right], \tag{6.57}$$

as a condition for α to deviate from 1 by not more than 10%. The conclusion is that the critical region is unobservably small, and the behavior seen, e.g., in the numerical solution discussed in Sec. VI.A.2.a, is determined by nonuniversal, preasymptotic power laws. As long as γ_t has not grown appreciably from γ_t^0, D_s and D will behave like

$$D_s\sim t^{s_s} \tag{6.58a}$$

with

$$s_s=\alpha/\alpha', \tag{6.58b}$$

and

$$D\sim t^{s_h} \tag{6.59a}$$

with

$$s_h=(\alpha-1)/\alpha'. \tag{6.59b}$$

Here α and α' are given by Eqs. (6.39a) and (6.39b), with $\tilde{\gamma}_t$ replaced by $\tilde{\gamma}_t^0$. They therefore depend on γ_t^0. With decreasing t, γ_t starts to grow, and the effective exponents acquire a t dependence. They are now given as solutions of Eqs. (6.39a) and (6.39b) as written. For $d=3$ and $F_0^a=-0.778$, α, α', s_s, and s_h are shown in Fig. 32. We see that s_h approaches zero very slowly. In the region covered by the inset in Fig. 31 (t on the order of a

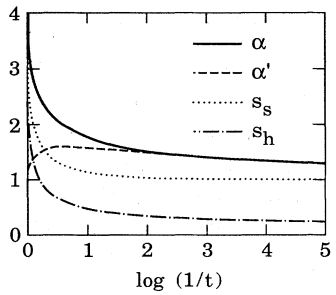

FIG. 32. Effective exponents α, α', s_s, and s_h as functions of $\ln(1/t)$ in the preasymptotic region. Parameters are $d=3$, and $F_0^a=-0.778$.

few percent), its value is close to 0.4. This is compatible with the numerical result. In contrast to s_h, s_s approaches its asymptotic value, 1, fairly rapidly. At $t=10^{-2}$, the deviation is less than 3%. This explains another feature of the numerical result, which had found D_s to vanish linearly (see Fig. 31).

In conclusion, the existence of the FL-IFS transition seems theoretically well established by a variety of techniques. However, the experimentally important question of how large a region in parameter space and IFS phase occupies has not been answered convincingly.

3. The metal-insulator transition

To describe the direct Fermi-liquid–insulator transition (Kirkpatrick and Belitz, 1992a) mentioned in the previous subsection we consider the two-loop RG equations derived in Secs. VI.A.1.a and VI.A.1.c,

$$b\frac{dg}{db}=-\varepsilon g+\frac{g^2}{4}[5-3(1+1/\gamma_t)\ln(1+\gamma_t)]\ , \tag{6.60a}$$

$$b\frac{d\gamma_t}{db}=\frac{g}{8}(1+\gamma_t)^2-\frac{3}{32}g^2\gamma_t^3[5/2-\ln(1+\gamma_t)]\ , \tag{6.60b}$$

$$b\frac{dh}{db}=\frac{h}{8}g(3\gamma_t-1)+\frac{3}{16}hg^2\gamma_t^2\ . \tag{6.60c}$$

Equations (6.60) contain the first two nonvanishing contributions to the flow equations for g, γ_t, and h. The one-loop terms in Eqs. (6.60) are exact. The two-loop terms were derived in the limit $\gamma_t\gg 1$. In order to describe a metal-insulator transition, one must use these equations for $\gamma_t\sim O(1)$ and $g\sim O(1)$. Though this procedure is, strictly speaking, uncontrolled, it can be justified *a posteriori* on grounds of the results' being qualitatively consistent with the existence of the multicritical point in Fig. 28 and with scenario (iv) for the metal-insulator transition discussed in Sec. IV.A.2.

The metal-insulator fixed point predicted by Eqs. (6.60) is defined by the equations

$$g^*=\frac{4\varepsilon}{5-3(1+1/\gamma_t^*)\ln(1+\gamma_t^*)}\ , \tag{6.61a}$$

$$5-3(1+1/\gamma_t^*)\ln(1+\gamma_t^*)=\frac{15}{2}\frac{\varepsilon\gamma_t^{*3}}{(1+\gamma_t^*)^2}[1-\tfrac{2}{5}\ln(1+\gamma_t^*)]\ . \tag{6.61b}$$

In $d=3$, the solution of Eqs. (6.61) is $g^*\simeq 4.02$ and $\gamma_t^*\simeq 0.84$. The linearized RG eigenvalues, which give the critical exponents for the metal-insulator phase transition, are determined by expanding Eqs. (6.60) about $(g^*,\gamma_t^*,h^*=0)$. One finds one relevant eigenvalue, $\lambda_+=1/\nu$, related to the correlation length exponent ν, and one irrelevant eigenvalue λ_-. This shows that the fixed point is stable. The linearization of the right-hand side of Eq. (6.60c) gives $h\kappa$, and the dynamical scaling exponent is $z=d+\kappa$ [cf. Sec. IV.A.2, scenario (iv)]. For $d=3$ the predicted exponents are

$$\nu=1/\lambda_+=0.75\ , \tag{6.62a}$$

$$\lambda_-=-4.08\ , \tag{6.62b}$$

$$z=5.91\ , \tag{6.62c}$$

$$\kappa=2.91\ . \tag{6.62d}$$

The density-of-states exponent β can be obtained by using the fixed-point values g^* and γ_t^* and Eqs. (6.6) and (6.7a). With $l_s\to-\pi/\sin(\pi\varepsilon/2)$, which is the generalization of $l_s=-2/\varepsilon$ to $\varepsilon=O(1)$, one finds

$$\beta=0.50\ . \tag{6.62e}$$

Near this metal-insulator transition the electrical conductivity σ, the spin susceptibility χ_s, and the coefficient of the linear term in the specific heat, $\gamma=\lim_{T\to 0}C/T$, satisfy the scaling laws given in Sec. IV.A. An approximate expression for the Fermi-liquid–insulator phase boundary in Fig. 28 can be determined by computing the RG critical surface for fixed point. In $d=3$,

$$g_0=3.38+0.78\gamma_t^0\ . \tag{6.63}$$

An important question is whether the fixed point discussed above is accessible. This has been addressed (Kirkpatrick and Belitz, 1992a) by means of a numerical solution of the flow equations, Eqs. (6.60). The result, shown in Fig. 33, was that there are several fixed points as a function of ε and initial values of g and γ_t. The physical fixed point, if any, is the one that first occurs with increasing disorder. For $\varepsilon=0$ [Fig. 33(a)] and $g_0=g(b=1)$ very small, g first increases and then decreases, and γ_t diverges at a finite value of b. The ground state in $d=2$ is therefore not a Fermi liquid. For larger g_0 one finds a stable fixed point $(g^*,\gamma_t^*)=(0.51,2.08)$ with an associated separatrix, separating regions in parameter space where γ_t diverges at a finite scale from those where g diverges at a finite scale. It is important to remember that Eqs. (6.60) are capable only of describing how metallic (small-g) behavior breaks down. For $\varepsilon=0$, this breakdown occurs already at arbitrarily small g_0 by means of $\gamma_t\to\infty$. The fixed point is therefore inaccessible and has no physical

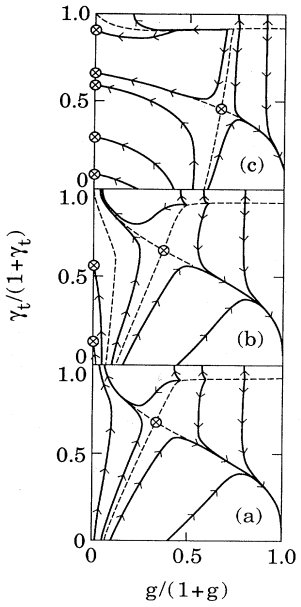

FIG. 33. Flow diagram showing the solution of Eqs. (6.60) for three different dimensionalities $d=2+\varepsilon$: (a) $\varepsilon=0$, (b) $\varepsilon=0.05$, (c) $\varepsilon=1$. Fixed points are denoted by crossed circles, and separatrices by dashed lines. From Kirkpatrick and Belitz (1992a).

significance. For $0<\varepsilon\ll 1$ [Fig. 33(b)] and small g_0, a Fermi-liquid fixed point ($g^*=0$, $\gamma_t^*=$finite number) is reached. With increasing g_0, one enters a region in which γ_t again diverges at a finite scale. The separatrix between these two regions can be related to the FL-IFS phase transition. Since it involves runaway trajectories, this transition requires an infinite resummation of the loop expansion and cannot be described by Eqs. (6.60). This has been discussed in Sec. VI.A.2 above. For still larger g_0, there is again an inaccessible stable fixed point.

For $\varepsilon=O(1)$, the behavior of the flow changes again qualitatively [Fig. 33(c)]. For small g_0 (and a γ_t^0 that is not too large), there always is a Fermi-liquid fixed point. With increasing g_0 one reaches the stable fixed point separating $g\to 0$ flow from $g\to\infty$ flow with no $\gamma_t\to\infty$ trajectories in between. The fixed point is therefore accessible and describes a metal-insulator transition (the FL-I phase boundary in Fig. 28). For $\gamma_t^0\gg 1$ there is again a separatrix related to the FL-IFS phase transition (the FL-IFS phase boundary in Fig. 28), confirming the existence of the multicritical point M in Fig. 28.

We conclude that for ε larger than a critical value $0<\varepsilon_c<1$ there is an accessible, stable fixed point describing a metal-insulator transition. Because there are no runaway trajectories at this point, one does not have to perform an infinite resummation as for the pseudomagnetic transition, but rather Eqs. (6.60) should be sufficient to describe it approximately. Furthermore, the results are consistent with both the results of Sec. VI.A.2 and scenario (iv) of Sec. IV.A.2, which makes it rather unlikely that the metal-insulator fixed point discussed above is an artifact of the two-loop RG equations. It must be stressed again, however, that the existing theory of this metal-insulator transition is uncontrolled. For the most part this is due to the fact that the transition simply does not exist for $d\to 2$. A controlled theory would therefore have to follow a route completely different from the one described here.

B. Experiments

The best studied material near a metal-insulator transition is phosphorus-doped silicon (Si:P), and some of its general features have already been mentioned in Sec. I. At the same time, it is the only material for which it has been established that it represents the generic universality class G at all practical temperatures. Since $\Delta Z=1$ in this material, the spin-orbit scattering is expected to be insignificant, and measurements of the spin-flip scattering rate (Paalanen *et al.*, 1986) have suggested that the spin-triplet channel survives down to $T\cong 3$ mK. In the millidegrees Kelvin temperature range, values of $t\simeq 10^{-3}$ have been achieved by using a stress tuning technique (Paalanen *et al.*, 1982). In this region both the transport and thermodynamics properties of Si:P near the metal-insulator transition have been measured very accurately. For these reasons we discuss only Si:P in this subsection. Also, consistent with the theory we have discussed, we concentrate on experimental results on the metallic side of the metal-insulator transition. In Sec. IX we shall discuss further aspects of the insulating phase.

1. Transport properties

In Fig. 34 the zero-temperature conductivity versus the dimensionless distance from the metal-insulator tran-

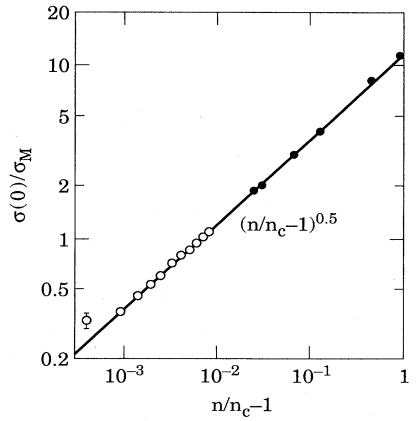

FIG. 34. Zero-temperature conductivity σ of Si:P, normalized by the Mott number, vs the dimensionless distance from the critical point. The solid line represents the best power-law fit. After Rosenbaum *et al.* (1983).

sition, $t=n/n_c-1$, is shown (Rosenbaum *et al.*, 1983). Here n is the phosphorous concentration, and $n_c \simeq 3.7\times10^{18}$ cm^{-3} is the critical density for the transition. The open circles in this figure were obtained using the stress tuning of a single sample, while the solid circles represent different samples. The solid line is a best fit assuming a power-law form: $\sigma=\sigma_0(n/n_c-1)^s$. The fit yields $s=0.51\pm0.05$. As emphasized in Secs. V.A.3.b and V.B, s obtained in this way should be interpreted as an effective exponent.

In Figs. 35 and 36 the temperature dependence of the conductivity is shown (Thomas *et al.*, 1983). In Fig. 35, the solid lines are linear fits to the data in the low-temperature limit. The dotted line is $\sigma(n_c,T)$, which is well represented by

$$\sigma(n_c,T)=(n/k_F)^2(T/T_F)^{0.37}\ , \tag{6.64}$$

with $T_F\cong100$ K the Fermi temperature. Taken at face value, this result gives an effective dynamical exponent equal to $z=1/0.37\simeq2.7$. However, the temperature range fit by this power law is very small. Figure 36 shows the slope of the temperature corrections to σ at low T. Note that $d\sigma/dT$ changes sign about 3% away from the metal-insulator transition. This is another example of the correlation noted in Sec. V.B between values of s that are smaller than $\frac{2}{3}$ and a sign change in the slope of $\sigma(T)$. As for the Si:B data discussed in Sec. V.B, we test whether or not the experimental data for Si:P actually satisfy scaling. In Fig. 37 we plot σ/t^s versus $x=T/t^\phi$ with $s=0.51$ and $\phi=sz=1.38$. The scatter in the experimental points indicates that scaling is *not* very well satisfied with these exponent values.

Very recently the interpretation of the data by Rosenbaum *et al.* has been questioned (Stupp *et al.*, 1994a; see also Rosenbaum, Thomas, and Paalonen, 1994, and Stupp *et al.*, 1994b). These authors assert that the rounding in the data ascribed by Rosenbaum *et al.* to sample inhomogeneities instead signals a crossover, and that the correct conductivity exponent is $s=1.3$.

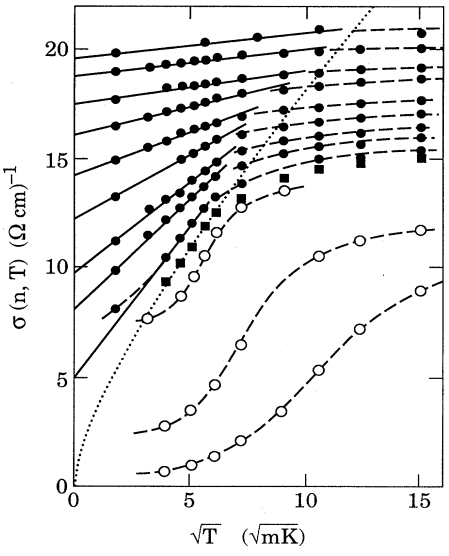

FIG. 35. Conductivity σ of Si:P as a function of temperature. The data fit by the dotted line represent a sample at the critical density. From Thomas *et al.* (1983).

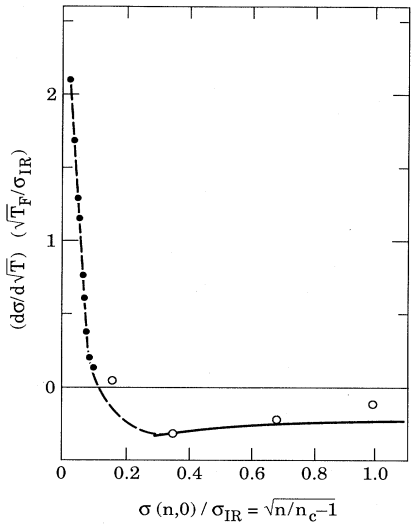

FIG. 36. Slope of the curves shown in Fig. 35, $d\sigma/d\sqrt{T}$, vs the dimensionless distance from the critical point. Note the change of sign. After Thomas *et al.* (1983).

We now illustrate how Eq. (5.44b) can be used to reconcile the experimental result for s with Eq. (4.16b) (Kirkpatrick and Belitz, 1993), assuming that the value for s obtained by Rosenbaum *et al.* is the correct one. If the effective exponent should indeed be 1.3, logarithmic corrections to scaling might still be present, but in the absence of a theoretical value for s in $d=3$ this experiment would not be a suitable test for their presence. Figure 38 shows some of the data of Fig. 34 together with

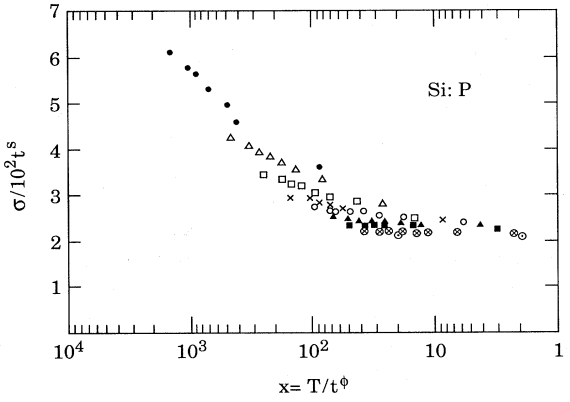

FIG. 37. Scaling plot of the conductivity data for Si:P. The assumed exponent values are $s=0.51$ and $\phi=sz=1.38$. The different symbols denote data taken at a given disorder t for various temperatures: ●, 6.04×10^{-4}; △, 1.40×10^{-3}; □, 2.20×10^{-3}; ×, 3.23×10^{-3}; ○, 4.22×10^{-3}; ▲, 5.52×10^{-3}; ■, 6.81×10^{-3}; ⊗, 8.50×10^{-3}; ⊙, 9.74×10^{-3}. The various temperatures were in the range 3 mK $< T <$ 53 mK.

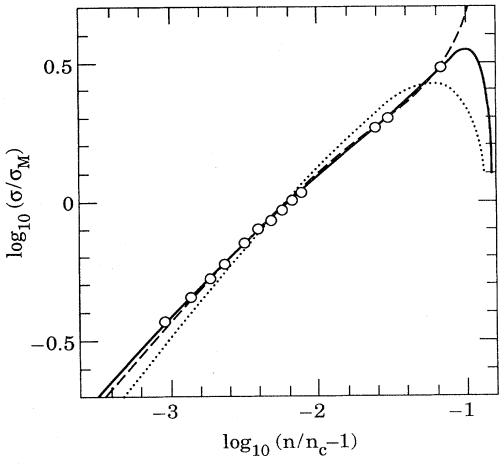

FIG. 38. Zero-temperature conductivity in Si:P vs dimensionless distance from the critical point. The dotted, dashed, and solid lines are best fits obtained by using one, two, and three correction terms in Eq. (5.42b) with an asymptotic critical exponent $s=0.7$. Best-fit values for the coefficients in Eq. (5.42b) are $\sigma_0=54.87$, 94.82, 132.16; $a_1=-1.84$, -4.42, -6.61; $a_2=0$, 6.21, 17.73; $a_3=0$, 0, -16.38 for the three curves, respectively. From Kirkpatrick and Belitz (1993).

best fits obtained from Eq. (5.44b) with $s=0.7$. The data points were selected as follows. For small t, roundoff due to sample inhomogeneities sets a limit at $t\cong 10^{-3}$. At large t, points up to $t=10^{-1}$ were included. In order to improve the statistics with several logarithmic correction terms taken into account, the 13 data points for $10^{-3}\le t<10^{-1}$ were augmented by another 12 obtained by linearly interpolating between neighboring points. A standard χ^2 fitting routine with singular-value decomposition was then used to optimize the values of the coefficients a_1, a_2, etc., in Eq. (5.44b). The dotted, dashed, and solid lines, respectively, in Fig. 38 represent the best fits obtained with one, two, and three correction terms taken into account. The fits obtained with two and three correction terms are of significantly higher quality than a straight-line fit optimizing s. More than three correction terms did not lead to further improvements in the fit quality. While the value for s was chosen arbitrarily, this shows that the experiment is certainly consistent with the theoretical lower bound for s, once corrections to scaling are taken into account. An attempt has also been made to determine s from the data by means of Eq. (5.44b), by repeating the fitting procedure with different values of s. However, the fit quality was found to go through a very shallow maximum as a function of s, with relatively large (± 0.15) fluctuations in the best value of s if one successively eliminated large-t data points. The best estimate for this experiment was $s=0.70^{+0.20}_{-0.03}$, where the lower bound was set by the theoretical bound rather than by the fit.

2. Thermodynamic properties

Figure 39 shows the donor susceptibility χ_D in Si:P normalized by the Curie susceptibility at $T=1$ K (Ootuka and Matsunaga, 1990). The diamagnetic contribution included in χ_D is not believed to lead to substantial deviations of χ_D from the spin susceptibility χ_s. As in the case of Si:P, B, Fig. 25, no saturation of χ_s is observed down to $T\cong 30$ mK, and the data are consistent with χ_s diverging for $T\to 0$ even well into the metallic phase. Also, the susceptibility shows no obvious additional anomaly across the metal-insulator transition. A power-law fit to the low-temperature behavior, $\chi_s\sim T^{-a}$, gives $a=0.46$. Similar behavior was observed by Paalanen *et al.* (1986).

The specific heat C in Si:P is shown in Fig. 40 as a function of temperature for three samples with $n/n_c=0.78$, 1.09, and 1.25 (Paalanen *et al.*, 1988). The phonon contribution, proportional to T^3, is shown as dashed lines for each of the three samples. Thereby a Debye temperature of 640 K has been assumed. The solid line represents the "free"-electron contribution $C_0=\gamma_0 T$, which was calculated using the Si conduction-band mass ($m_0^*=0.34m_0$), consistent with specific-heat measurements well above n_c. In general the anomaly in the metallic phase is not strong enough to be meaningfully characterized by an exponent. For the insulating sample ($n/n_c=0.78$), the specific heat is consistent with a power law $\gamma/\gamma_0\sim T^{-0.6}$ below 0.7 K. For completeness we show in Fig. 41 the specific heat well into the insulating phase (Lakner and v. Löhneysen, 1989). An interesting feature is that, for small enough n, the specific heat increases with decreasing temperature down to the mK temperature range. Since $C(T\to 0)$ must vanish, this implies that, at least in the insulating phase of Si:P, there is

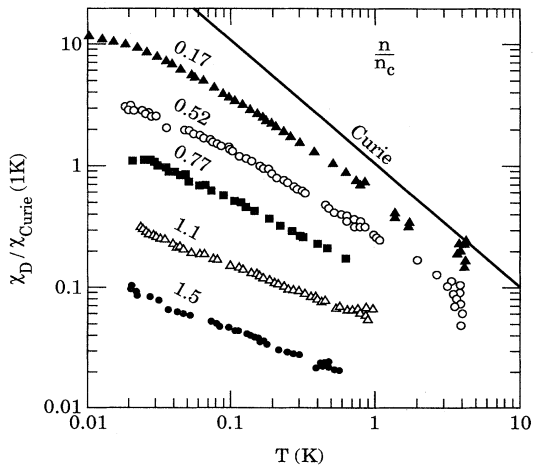

FIG. 39. Donor susceptibility χ_D in Si:P, normalized to the Curie susceptibility at 1 K, vs temperature. The labeling of the curves gives the phosphorus concentration n/n_c, and the solid line represents the Curie susceptibility. From Ootuka and Matsunaga (1990).

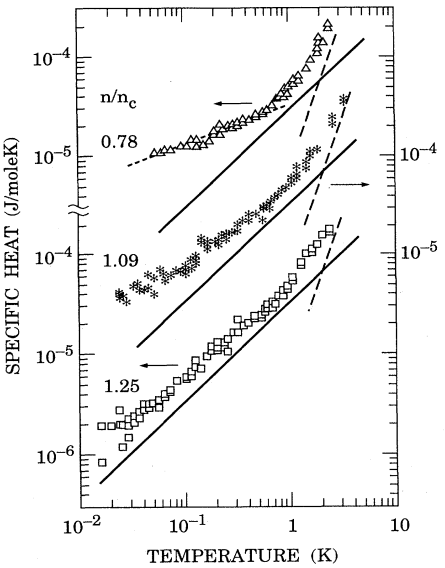

FIG. 40. Specific heat of Si:P as a function of temperature for three different phosphorus concentrations. See text for more information. After Paalanen *et al.* (1988).

an extremely low temperature scale that is not understood.

The thermodynamic data shown in Figs. 39–41 are not readily explained with the theoretical results of Secs. V.A and VI.A. Because Si:P is expected to be in the generic universality class, it is in principle possible that the IFS phase exists between the Fermi liquid and the insulator (cf. Fig. 28) and that this is what is observed in some of these experiments. However, the best available estimate for the location of the multicritical point is $|F_0^a|^c \gtrsim 0.7$, while a reasonable value for Si:P is $F_0^a \cong -0.5$. It is therefore more likely that the experiments probe the FL-I transition than that they probe the FL-IFS phase transition. Even more importantly, in the Fermi-liquid phase, on the low-disorder (or high-doping in the case of Si:P) side of the FL-IFS phase transition, one expects χ_s to saturate at low temperatures. No such saturation has been observed, but since no theoretical estimate for the saturation temperature is available it is possible that the experimental temperatures have not been low enough. An alternative explanation of the data in terms of local magnetic moments in both the metallic and insulator phases will be discussed in Sec. IX.

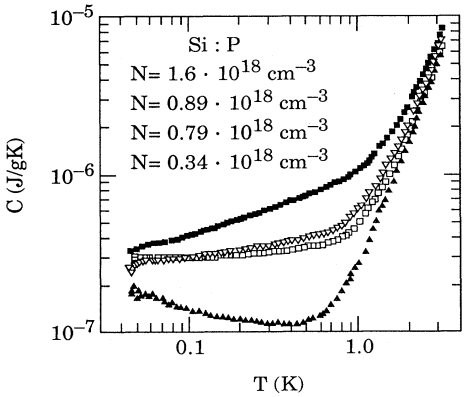

FIG. 41. Specific heat C vs temperature T of Si:P for various donor densities. The critical density is about 3.3×10^{18} cm^{-3}. From Lakner and von Löhneysen (1989).

VII. DESTRUCTION OF CONVENTIONAL SUPERCONDUCTIVITY BY DISORDER

In this section we consider the effects of nonmagnetic disorder on systems whose ground state in the clean limit is a superconductor rather than a Fermi liquid. This is a rather extensive subject, and complete coverage of all aspects would require a separate review. A further complication arises from the fact that the phenomenology of the field is complex and not as widely known as that of the metal-insulator transition. Also, to the authors' knowledge, the topic has never been reviewed in full, though aspects of it have been covered by Bergmann (1976), Rowell and Dynes (1980), Lee and Ramakrishnan (1985), Ramakrishnan (1989), and Belitz (1990). We therefore give a brief overview of the field in Sec. VII.A, where we also explain the limitations of our coverage. In Sec. VII.B.1 we describe the results, but not the detailed derivations, of various generalizations of the standard theories of superconductivity to disordered systems. Section VII.B.2 contains a detailed account of recent approaches that use the field-theoretic methods that are the main theme of this review, and Sec. VII.C discusses experiments. We assume that the reader is familiar with the Ginzburg-Landau, BCS, and Eliashberg theories [see the books by Abrikosov *et al.* (1975), Bogoliubov *et al.* (1959), de Gennes (1966), Parks (1969), Schrieffer (1964), or Tinkham (1975)].

A. Disorder and superconductivity: A brief overview

Reports of the influence of disorder on superconductivity date back to Kamerlingh Onnes in 1925, but the first systematic study of various materials was performed by Buckel and Hilsch (1954). By quench-condensing thin films of superconducting elements on cold substrates they obtained residual resistivities of typically 20 $\mu\Omega$ cm. For Al, Zn, Sn, In, and Tl the transition temperature T_c was found to be enhanced compared to the clean bulk value, with enhancement factors ranging from 2.3 (Al) to 1.1 (Tl). The T_c of Pb was found not to change, and that of Hg to decrease slightly. Matthias *et al.* (1956) studied solid solutions with up to 50% of chemical scattering centers. In cases where the electron concentration was not expected to change much in the alloying process, they found T_c values comparable to those of the pure materials. Lynton *et al.* (1957) systematically studied the effect of very small ($\sim$0.1 at. %) amounts of impurities on the T_c of Sn, starting from single crystals. They found an initial sharp drop in T_c followed by a much more gradual behavior at larger impurity concentrations. In contrast to these mild effects of substantial amounts of

nonmagnetic scattering centers, magnetic impurities were found to lead to a continuation of the initial sharp drop of T_c, and T_c vanished at concentrations on the order of 1 at. % (Matthias *et al.*, 1958).

This qualitative difference between the effects of magnetic and nonmagnetic disorder was explained on the basis of BCS theory (Bardeen *et al.*, 1957). Abrikosov and Gorkov (1960) showed how magnetic impurity scattering breaks the spin-singlet Cooper pairs and suppresses T_c. For nonmagnetic disorder, Abrikosov and Gorkov (1958, 1959) showed that, to leading order in perturbation theory, T_c is independent of disorder; see Sec. III.B.2.b. Anderson (1959b) gave a qualitative explanation of why nonmagnetic disorder should have no drastic effects. He considered the set of exact eigenstates, ψ_α, of a Hamiltonian describing noninteracting electrons in a disordered environment, Eq. (2.1a). While the ψ_α cannot be determined explicitly, they form a complete set and can be used as a basis set analogous to plane waves in clean materials. In particular, one can form a Cooper pair out of a ψ_α and its time-reversed counterpart. BCS theory can then be repeated with the ψ_α replacing plane waves, and the only effects of disorder are small changes in the matrix elements and in the density of states upon ensemble averaging. These changes are expected to be largest in the initial stages of the disordering process, where peaks in the density of states and in the matrix elements get smeared out. This accounts for the observed initial sharp drop in T_c. Once this has happened, one would expect no further change of T_c with increasing disorder. This prediction is sometimes called "Anderson's theorem." The dependence of T_c on the single-particle DOS that remains even within the exact eigenstate formulation of the BCS theory has been invoked to explain not only the small changes of T_c at small disorder mentioned above, but also the large changes that occur over a substantial range of disorder in, e.g., the A-15 materials (Testardi and Mattheiss, 1978). This kind of one-electron description was criticized by Anderson *et al.* (1983), and it has become increasingly clear that the experimental observations suggest a breakdown of Anderson's theorem due to an interplay between disorder and many-body effects.

It has always been clear that Anderson's theorem cannot be strictly valid for real materials (though it is strictly valid within certain models; see below), but our understanding of how and why T_c depends on nonmagnetic disorder has progressed slowly. Much experimental information has accumulated over the years which shows that, though small amounts of disorder can either increase or decrease T_c, sufficiently large disorder[18] will always destroy superconductivity. The subject gained popularity among theorists in the 1980's, when the interest in localization raised the question of whether the destruction of superconductivity is due to incipient or actual localization of one-electron states. A particularly interesting case is that of 2-*d* systems, whose ground state is insulating in the absence of superconductivity. This gave rise to the somewhat ill-posed question of whether or not superconductivity is possible in systems where the one-particle excitations are localized.

In bulk systems, the observed degradation of T_c at sufficiently large disorder has often been found to depend only on the extrapolated residual resistivity, not on how the disorder was introduced. This has at times given rise to the notion of a "universal" T_c degradation. Recently, however, through systematic studies on thin films, it has become clear that the morphology of the materials is very important for their superconducting properties. From a theoretical point of view it should be stressed that there is no universality, in the technical sense of the word, in the T_c degradation. It depends on many nonuniversal details of the system, just as the value of T_c in the clean limit does. It has proven useful to consider two idealized classes of materials, viz., "homogeneous" and "granular" disordered superconductors (Deutscher *et al.*, 1985; Valles and Dynes, 1990). The homogeneous systems are thought to be disordered on an atomic length scale as one would expect, e.g., in a solid solution or as a result of irradiation. The granular systems, on the other hand, are composed of grains or microcrystallites of relatively clean material with linear dimensions of typically 50 Å, which are separated by regions of normal conducting or insulating material. The prototype of a "homogeneous" system is, apart from bulk solid solutions, a Pb film grown on a substrate with an underlayer of Ge (Strongin *et al.*, 1970); the prototype of a granular system is granular aluminum (see, for example, Abeles, 1976). The former show a steady T degradation, but the transition remains sharp (Fig. 42), and superconductivity is believed to be destroyed by a suppression of the *amplitude* of the superconducting order parameter. In the latter, T_c defined as the onset of the transition stays almost constant while the transition becomes steadily broader (Fig. 43), which is believed to be due to increasing fluctuations of the *phase* of the order parameter (Valles *et al.*, 1989; Valles and Dynes, 1990). Real systems often lie in between these extremes, and often it is not known how "homogeneous" a given sample really is.

In this section we cannot possibly cover all aspects of disordered superconductors. In accordance with the general theme of this review, we concentrate on the destruction of superconductivity which happens more or less close to the metal-insulator transition. We restrict our discussion to the disorder dependence of T_c and do not consider effects in the superconducting state. We further restrict our discussion to homogeneous systems, where the relevant physics is believed to be close to that governing the other topics in this review, namely, interplay of

[18]What "sufficiently large" means depends on the material, and in particular on the value of T_c in the clean limit, T_c^0. Systems with a low T_c^0 almost always show an initial increase of T_c upon disordering, while those with a high T_c^0 show an immediate T_c degradation. This correlation has been noted by Belitz (1987a).

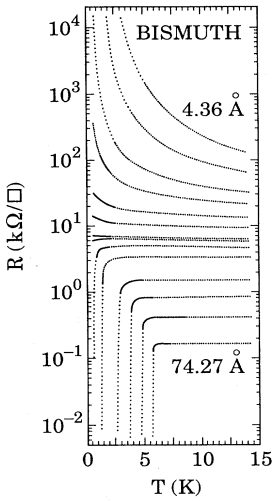

FIG. 42. Resistance as a function of temperature for a series of homogeneous Bi films. From Haviland *et al.* 1989).

disorder and interactions. A brief overview including the granular models as well has been given by Ramakrishnan (1989). For homogeneous thin films, it has been suggested that there is a *universal* normal-state sheet resistance $R_{\square}^{*} \cong h/4e^2$, which separates superconducting samples from insulating ones (Jaeger *et al.*, 1986; Haviland *et al.*, 1989). There is currently no agreement between experimental groups as to whether or not this is actually the case. We shall briefly comment on a theoretical argument in favor of a universal $R_{\square}^{*}$ in Sec. IX.D. In this section we also exclude effects in magnetic fields, and we restrict ourselves to the BCS T_c, even in two dimensions, where strictly speaking there is no BCS transition. We deal only with conventional, phonon-mediated, *s*-wave superconductivity. We do not discuss heavy-fermion or high-T_c superconductors (a discussion of disorder effects in heavy-fermion superconductivity can be found in Ahlheim *et al.*, 1988).

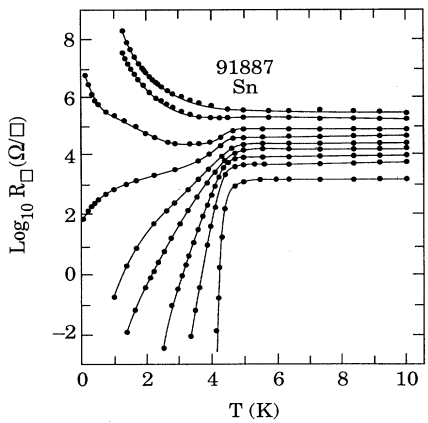

FIG. 43. Resistance as a function of temperature for a series of granular Sn films. From Valles and Dynes (1990). The lines are guides to the eye.

B. Theories for homogeneous disordered superconductors

1. Generalizations of BCS and Eliashberg theory

In homogeneous systems, the destruction of superconductivity is preceded by a characteristic degradation of T_c. It is possible, at least in $d=3$, that essentially the same physics that is responsible for this degradation also brings about the eventual destruction of superconductivity. Before we discuss methods capable of describing the disappearance of superconductivity, we therefore mention the substantial body of work aimed at describing this degradation by adding disorder, often perturbatively, to the standard theories of superconductivity.

a. Enhancement and degradation of the mean-field T_c

Keck and Schmid (1975, 1976) considered Eliashberg theory with an electron-phonon coupling given by Eqs. (2.46). They dressed the bare vertex by means of impurity ladder diagrams as explained in Sec. II.A.2.b and calculated the anomalous self-energy in a conserving approximation in the sense of Kadanoff and Baym (1962). The result was the linearized Eliashberg (1960) equations in standard form with all of the disorder dependence in a generalized Eliashberg function,

$$\alpha^2 F(\omega) = \alpha^2 F_L(\omega) + 2\alpha^2 F_T(\omega) \ , \tag{7.1a}$$

with

$$\alpha^2 F_b(\omega) = \frac{6 N_F d_b}{\pi k_F^2 l} \int_0^\infty dq \, B_b(q,\omega) f_b(ql) \ . \tag{7.1b}$$

Here $b = L, T$ for longitudinal and transverse phonons, respectively, l is the electronic mean free path, $B_b(q,\omega)$ is the phonon spectral function, and d_b and f_b are given by Eqs. (2.50). The disorder dependence of T_c is then given through that of the usual coupling constant

$$\lambda = 2 \int d\omega \, \alpha^2 F(\omega)/\omega \ . \tag{7.2}$$

Since there is no diffusion enhancement of the electron-phonon vertex [see Eq. (2.48)], this disorder dependence is weak. Furthermore, it comes from coupling to phonons with wave number $q \cong 1/l$ and therefore is model dependent. Keck and Schmid evaluated $T_c(l)$ for a Debye model and found a decrease in the effective coupling to longitudinal phonons, which is overcome by an increase (from zero) in the coupling to transverse phonons, so that a net increase of T_c linear in $1/k_F l$ results. They found this increase to be consistent with experiments on In by Bergmann (1969).

The work of Keck and Schmid has often been criticized, both on theoretical and experimental grounds. As mentioned in Sec. II, many other calculations (incorrectly) obtained a diffusion-enhanced electron-phonon vertex (see, e.g., the discussion by Schmid, 1985), which leads to

an increased coupling to *longitudinal* phonons. Some of these theories did not use the Tsuneto-Schmid transformation to a moving coordinate system, but rather considered the effects of moving impurities in a stationary frame of reference. Since many workers considered the latter method more convincing, this cast doubt on Schmid's calculation. The resulting argument has been put to rest by Reizer and Sergeyev (1986). They showed that a correct calculation with either method yields for α^2F the result of Keck and Schmid, Eq. (7.1). They also pointed out, however, that care must be taken in drawing conclusions about T_c, which depends on the electron-phonon coupling at large wave numbers as well as small ones. Criticism from the experimental side arose from the fact that Eq. (7.1) with Debye phonons gives a low-frequency asymptotic behavior $\alpha^2F(\omega\rightarrow 0)\propto\omega^3$, while most experiments find a linear low-frequency behavior (Bergmann, 1976). This discrepancy has been discussed by Belitz (1987b), who showed that the inclusion of phonon damping in Eq. (7.1b) qualitatively changes the low-frequency asymptotics. As a result, α^2F shows a linear frequency dependence both at asymptotically low frequencies and at larger frequencies $\omega\gg c/l$ with c the speed of sound. The ω^3 law appears at most in an intermediate-frequency window whose width is disorder dependent, and which is absent for resistivities smaller than about 50 $\mu\Omega$ cm.

Since Keck and Schmid did not consider the Coulomb repulsion, their theory predicts that T_c will *always* increase with disorder. Belitz (1987a) has argued that for low-T_c materials at small disorder the Keck-Schmid mechanism (with some modifications due to crossed-ladder renormalizations, see below) indeed dominates all T_c degradation effects and may well explain the experimentally observed T_c enhancement in this regime. This theory has recently been found to be in very good agreement with new experimental investigations by Miehle *et al.* (1992).

Disorder-induced degradation of T_c was first considered by Maekawa and Fukuyama (1981, 1982) and by Takagi and Kuroda (1982). These authors considered 2-d systems within a BCS model, where λ is a phenomenological constant. Their perturbative disorder renormalizations of the pair propagator and the Coulomb pseudopotential yielded a T_c decrease linear in $1/k_Fl$. If the theory is extrapolated to the regime of small T_c, it predicts reentrant behavior, i.e., the ground state is predicted to be an insulator. Reentrance has not been observed experimentally, but the prediction is clearly outside of the region of validity of the perturbation theory. The initial decrease in T_c explains the experimentally observed behavior with reasonable parameters (e.g, Graybeal and Beasley, 1984). In $d=3$, the same theory (Fukuyama *et al.*, 1984; Maekawa *et al.*, 1984) yielded a leading correction to T_c which is proportional to $1/(k_Fl)^2$. This result has been criticized by Belitz (1985), who argued that the correct result is of order $1/k_Fl$ and that the extra factor of $1/k_Fl$ is due to the same mistake as was discussed in connection with Eq. (2.41).

Anderson *et al.* (1983) considered the local Coulomb kernel in a strong-coupling theory (Scalapino *et al.*, 1966),

$$K^c(i\omega_n)=\sum_{\mathbf{q}}K^c(\mathbf{q},i\omega_n)=\sum_{\mathbf{q}}V(\mathbf{q},i\omega_n)\Phi_{\rho\rho}(\mathbf{q},i\omega_n)\ , \tag{7.3}$$

where V is the dynamically screened Coulomb potential, Eq. (2.45a), and $\Phi_{\rho\rho}$ is the density-density Kubo function for noninteracting electrons, Eq. (2.36) [for a derivation of Eq. (7.3), see Gutfreund *et al.* 1985]. The local Coulomb kernel is a frequency-dependent generalization of the instantaneous Coulomb potential μ in the clean limit. In the Coulomb pseudopotential μ^* (Morel and Anderson, 1962), μ appears both in the numerator and in the denominator, so a mere increase of μ would have little effect. Anderson *et al.* noted, however, that for sufficiently strong disorder K^c develops a strong retardation, with K^c at low frequencies much more strongly enhanced than at high frequencies. They then solved the linearized gap equation using a two-square well approximation and obtained for T_c a McMillan formula with a generalized Coulomb pseudopotential of the form

$$\mu^*=\mu'/[1+\mu'\ln(\varepsilon_F/\omega_D)-(\mu'-\mu)\ln(\varepsilon_F\tau\sqrt{\alpha})]\ , \tag{7.4a}$$

where

$$\mu'=\mu[1+(9\pi/4k_F^2l^2)\ln\alpha] \tag{7.4b}$$

and

$$\alpha=(\rho/\rho_c)^3\ , \tag{7.4c}$$

with ρ the residual resistivity and ρ_c the characteristic resistivity scale at the Anderson transition. For systems sufficiently close to the Anderson transition, μ^* is enhanced and T_c is suppressed accordingly. Comparison with the observed T_c degradation in A-15 materials (Rowell and Dynes, 1980) yielded good agreement if ρ_c was taken in the range 7–31 $\mu\Omega$ cm, about a factor of 50 smaller than one would expect. While Anderson *et al.* gave some reasons for why ρ_c might be small in A-15's, Gutfreund *et al.* (1985; Entin-Wohlman *et al.*, 1986) argued that these materials with resistivities in the 100 $\mu\Omega$ cm range cannot possibly be close to an Anderson transition. They proposed instead that the quasi-one-dimensional nature of these materials is important. Accordingly, they assumed a very anisotropic $\mathbf{q}$ dependence of the integrand in Eq. (7.3). This strongly enhances the increase of μ^*, but the authors concluded that still the effect by itself was not strong enough to explain the T_c degradation in A-15's with realistic parameters. Leavens (1985) then found that an inclusion of the critical retardation effects of Anderson *et al.* in the normal self-energy function Z, which these authors had taken equal to $1+\lambda$ as in the clean case, actually leads to an *increase* of T_c close to the Anderson transition. This is contrary to what is observed in experiment and makes it clear that

there must be other T_c-degrading effects.

The work of Wysokinski and Kuzemski (1983) implicitly contains both of the effects discussed so far, albeit for a different model. The retardation and enhancement of the Coulomb kernel, Eq. (7.3), is implicit in their effective Coulomb potential $U_{\rm eff}$, but was not explicitly discussed.

Belitz (1987c, 1987d) has constructed a strong-coupling theory that contains all of the above effects and two additional ones. First, the theory considers crossed-ladder renormalizations in addition to the usual impurity ladders. Secondly, it was found that, besides the usual renormalization function Z, an additional normal self-energy piece is important in disordered systems. This contribution, which was denoted by Y, is even in frequency, and in the clean case it is an uninteresting constant, which simply renormalizes the Fermi energy (Eliashberg, 1960). In the disordered case it reflects the effects of the Coulomb gap (Altshuler and Aronov, 1979a, 1979b) on T_c and has to be kept. The importance of Y was noted independently by Browne *et al.* (1987). After several approximations, the theory could be cast into the same structure as Eliashberg theory and was solved by McMillan's method. The resulting T_c formula is

$$T_c=\frac{\Theta_D}{1.45}\exp\left[\frac{-1.04(1+\tilde{\lambda}+Y')}{\tilde{\lambda}-\mu^*[1+0.62\lambda/(1+Y')]}\right]. \tag{7.5}$$

Here $\tilde{\lambda}$, Y' (which is related to the self-energy piece Y), and μ^* are all monotonically increasing functions of disorder. $\tilde{\lambda}$ contains the Keck and Schmid effect, and μ^* that of Anderson *et al.* With this result, reasonable fits to experiments on A-15's and on rare-earth rhodium borides have been obtained.

b. Breakdown of mean-field theory

The reason why mean-field theory works so well for superconductivity is the large value of the coherence length ξ_0. In a disordered system, i.e., one with diffusive quasiparticle dynamics, the coherence length is $\xi=(\xi_0 l)^{1/2}$. With decreasing mean free path l, ξ decreases, which according to the Ginzburg criterion increases the size of the region where thermal fluctuations are important. If superconductivity survives until a region of sufficiently short ξ is reached, one would then expect fluctuations to play an important role in the suppression of superconductivity, even in homogeneous systems.

Kapitulnik and Kotliar (1985; Kotliar and Kapitulnik, 1986) have considered these questions in the framework of a Ginzburg-Landau theory. They concluded that in the critical region of the Anderson transition ξ saturates at a value proportional to $(N_F\Delta)^{-1/3}$, with Δ the superconducting gap and a prefactor of order unity. This gives rise to a sizable fluctuation region and reduces T_c below its mean-field value. Ma and Lee (1985) have argued that, within mean-field theory, superconductivity can survive even on the insulating side of the metal-insulator transition, a rather extreme example of Anderson's theorem in which localized one-particle states are paired. Ma *et al.* (1986) then mapped the problem onto a quantum spin-$\frac{1}{2}$ model with a random field. They found a superconductor-insulator transition induced by quantum fluctuations. Recently Smith and Ambegaokar (1992) have argued that coherent backscattering events decrease the number of superconducting electrons to the point where fluctuations become important. None of these theories considered the disorder enhancement of the Coulomb repulsion, which from perturbation theory is known to be quite strong (see the previous subsection). Their applicability to real superconductors is therefore not obvious.

2. Field-theoretic treatments

In this section we use the methods explained in Sec. III to describe disordered superconductors. Two distinct approaches will be discussed in detail. One is based on a microscopic derivation of a Landau-Ginzburg theory for superconductivity in disordered electronic systems. The net result is a Landau theory for the superconducting order parameter, with coefficients given in terms of correlation functions of a disordered electronic system in the absence of superconductivity. This result (Maekawa and Fukuyama, 1982) can be derived using a variety of techniques. To stay within the spirit of this review we shall give a field-theoretic derivation due to Kirkpatrick and Belitz (1992c). The resulting theory can then be used to study the suppression of the mean-field critical temperature by disorder. This approach should be expected to be subject to the usual drawbacks and advantages of any Landau theory: (1) It will work best in high spatial dimensions, but will be applied to $d=3$ and even to $d=2$ anyway. (2) The theory should give, at least in $d=3$, a qualitatively correct picture of the superconducting transition and of the physical mechanisms that determine the disorder dependence of T_c.

The other approach we shall discuss in detail is due to Finkel'stein (1987). It is based on the nonlinear sigma model given in Sec. III, with $K^{c(s)}<0$, so that in the absence of disorder the system is superconducting at low temperatures. Again, we shall discuss how to determine the disorder dependence of T_c using this approach.

We end this subsection with a discussion of how these two approaches are related to each other and to other work on T_c degradation. We also briefly discuss other recent descriptions of disordered superconductors.

a. The Landau-Ginzburg-Wilson functional

The main idea behind this approach is to separate the superconducting fluctuations in the particle-particle channel from the disorder-induced charge- and spin-density fluctuations in the particle-hole channel, which we discussed in Secs. V and VI. The net result is an

effective field theory for the superconducting order parameter. The Gaussian approximation to this field theory reproduces the mean-field theory of Maekawa and Fukuyama (1982). In spirit the approach is a generalization of early work by Gorkov (1959).

We start with the action for a general disordered, interacting Fermi system as derived in Sec. III.B.1,

$$S = S_0 + S_{\rm dis} + S_{\rm int}^{(s)} + S_{\rm int}^{(t)} + S_{\rm int}^{c(s)} \equiv \tilde{S}_0 + S_{\rm int}^{c(s)} \ . \tag{7.6}$$

S_0 and $S_{\rm dis}$ are given by Eqs. (3.43) and(3.44), respectively, and $S_{\rm int}^{(s,t)}$ by Eqs. (3.48b) and (3.48c). In the spin-singlet Cooper channel, we neglect the wave-number dependence of the interaction. Then Eq. (3.51b) yields

$$S_{\rm int}^{c(s)} = \frac{T}{2} \sum_{\sigma \neq \sigma'} \sum_{\mathbf{k},\mathbf{p}} \sum_{\mathbf{q}} \sum_{\substack{n_1 \\ m}} \sum_{n_2} \gamma^c_{n_1 n_2}(m) \bar{\psi}_\sigma(\mathbf{k}, n_1) \bar{\psi}_{\sigma'}(-\mathbf{k}+\mathbf{q}, -n_1+m) \psi_{\sigma'}(-\mathbf{p}+\mathbf{q}, -n_2+m) \psi_\sigma(\mathbf{p}, n_2), \tag{7.7}$$

where for simplicity we have suppressed the replica index. For the mean-field solution given here, this index does not play a role. For a more general treatment that includes fluctuation effects it can and must be retained. Further, since we are interested in singlet superconductivity, we put the spin-triplet Cooper channel interaction equal to zero. In Eq. (7.7), an attractive Cooper channel interaction means $\gamma^c > 0$. The partition function can then be written

$$Z = \tilde{Z}_0 \langle \exp(S_{\rm int}^{c(s)}) \rangle_{\tilde{S}_0} \ , \tag{7.8a}$$

where

$$\tilde{Z}_0 = \int D[\bar{\psi}, \psi] e^{\tilde{S}_0[\bar{\psi}, \psi]} \ , \tag{7.8b}$$

and the average is to be taken with the action $\tilde{S}_0$ defined by the second equality in Eq. (7.6),

$$\langle (\cdots) \rangle_{\tilde{S}_0} = (1/\tilde{Z}_0) \int D[\bar{\psi}, \psi] (\cdots) e^{\tilde{S}_0[\bar{\psi}, \psi]} \ . \tag{7.8c}$$

The idea is now to decouple $S_{\rm int}^{c(s)}$ by means of a Hubbard-Stratonovich transformation, but to leave the particle-hole channel in terms of the Grassmann fields. This is in contrast to Sec. III.B.1.c, where all four-fermion terms were decoupled and the Grassmann variables were integrated out.

We define a Grassmann field

$$\Psi_{\substack{n \\ \sigma\sigma'}}(\mathbf{q}, m) = \sum_{\mathbf{k}} \psi_\sigma(-\mathbf{k}+\mathbf{q}, n+m) \psi_{\sigma'}(\mathbf{k}, n) \tag{7.9}$$

and consider classical fields $Y_n(\mathbf{q}, m)$, $\bar{Y}_n(\mathbf{q}, m)$, and a metric

$$\bar{Y} \cdot Y = \sum_{n_1, n_2} \bar{Y}_{n_1} \gamma^c_{n_1 n_2} Y_{n_2} \ . \tag{7.10}$$

Then we can write the partition function in terms of a functional integral over Y,

$$Z = \int D[\bar{Y}, Y] e^{-\sum_q \bar{Y}(q) \cdot Y(q)} e^{F[\bar{Y}, Y]} \ , \tag{7.11a}$$

where $q = (\mathbf{q}, m)$ as in Sec. III, and

$$F[\bar{Y}, Y] = \ln \left\langle e^{-\sum_q [\bar{Y}(q) \cdot \Psi(q) + \bar{\Psi}(q) \cdot Y(q)]} \right\rangle_{\tilde{S}_0} \ . \tag{7.11b}$$

As in Sec. III, we omit multiplicative constant contributions to Z. We define the superconducting order parameter by scaling Y with γ^c:

$$\Delta_n(q) = \sum_m \gamma^c_{nm}(q) Y_m(q) \ . \tag{7.12}$$

Then we have

$$Z = \int D[\bar{\Delta}, \Delta] e^{-S_{\rm LGW}[\bar{\Delta}, \Delta]} \ , \tag{7.13a}$$

with a Landau-Ginzburg-Wilson action

$$S_{\rm LGW}[\bar{\Delta}, \Delta] = \sum_{\mathbf{q}} \sum_{n_1 n_2} \bar{\Delta}_{n_1}(\mathbf{q}) (\gamma^c)^{-1}_{n_1 n_2}(\mathbf{q}) \Delta_{n_2}(\mathbf{q}) - F[\bar{\Delta}, \Delta] \ , \tag{7.13b}$$

which can be expanded in powers of Δ,

$$S_{\rm LGW}[\bar{\Delta}, \Delta] = \sum_{\mathbf{q}} \sum_{m} \sum_{n_1 n_2} \bar{\Delta}_{n_1}(\mathbf{q}, m) [(\gamma^c)^{-1}_{n_1 n_2}(m) - \pi T C_{n_1 n_2}(\mathbf{q}, m)] \Delta_{n_2}(\mathbf{q}, m) + O(\bar{\Delta}^2 \Delta^2) \ . \tag{7.14}$$

Here C denotes the pair propagator,

$$C_{n_1 n_2}(\mathbf{q}, m) = \sum_{\mathbf{k},\mathbf{p}} \langle \psi_\uparrow(-\mathbf{k}+\mathbf{q}, -n_1+m) \bar{\psi}_\downarrow(\mathbf{k}, n_1) \psi_\downarrow(\mathbf{p}, n_2) \bar{\psi}_\uparrow(-\mathbf{p}+\mathbf{q}, -n_2+m) \rangle_{\tilde{S}_0} \ . \tag{7.15}$$

We see that the coefficients in the expansion (7.14) of $S_{\rm LGW}$ are given in terms of electronic correlation functions for a reference system that contains the full interaction in the particle-hole channel as well as the full disorder. The coefficient for the Gaussian term, Eq. (7.14), is a four-point correlation function, viz., the pair propagator. Coefficients for higher terms are determined by higher-order correlation functions. In Secs. V and VI we have calculated the most general four-point correlation function for a general interacting system. In particular, we

know the pair propagator for our reference system, and we shall now make use of this knowledge for a mean-field theory of disordered superconductors.

The Landau theory of the superconducting phase transition is given by the identification of $S_{\mathrm{LGW}}[\bar{\Delta},\Delta]$ with an effective potential for the superconducting state. The mean-field transition temperature is given by the extremum condition

$$\frac{\delta S_{\mathrm{LGW}}}{\delta \bar{\Delta}_n(q)} = \frac{\delta S_{\mathrm{LGW}}}{\delta \Delta_n(q)} = 0 \ . \tag{7.16}$$

Assuming that the transition is continuous, Eq. (7.16) leads to the linearized gap equation

$$\Delta_n(\mathbf{q},m) = \pi T \sum_{n_1 n_2} \gamma^c_{n n_1}(m) C_{n_1 n_2}(\mathbf{q},m) \Delta_{n_2}(\mathbf{q},m) \ , \tag{7.17}$$

which is an explicit equation for T_c. The superconducting instability occurs first for $\mathbf{q}=m=0$. We assume that γ^c has the form

$$\gamma^c_{n_1 n_2}(m=0) = |K^{c(s)}| \Theta(N-n_1)\Theta(N-n_2) \ . \tag{7.18}$$

Here $K^{c(s)} \equiv K^{(3)} < 0$ is the spin-singlet Cooper channel interaction amplitude from Eq. (3.92h), and $N = \omega_D / 2\pi T$, with ω_D a frequency cutoff on the order of the Debye frequency. With this ansatz the integral equation (7.17) is separable, and we get for the T_c equation

$$1/|K^{c(s)}| = \pi T \sum_{n_1 n_2}{}' C_{n_1 n_2}(\mathbf{q}=0, m=0) \ , \tag{7.19}$$

where the prime on the summation symbol indicates the ω_D cutoff.

Notice that the pair propagator C, Eq. (7.15), contains all interactions except for the BCS interaction in the particle-particle channel, Eq. (7.18). Equation (7.19) can be derived by a variety of methods (Maekawa and Fukuyama, 1982). The advantage of the above derivation (Kirkpatrick and Belitz, 1992c) is that it leads to an order-parameter field theory, so that it is obvious, in principle, how to go beyond mean-field theory. We shall discuss some aspects of superconducting fluctuations in the following subsections. Within mean-field theory the remaining problem is to calculate the pair propagator. This has been done in perturbation theory by Fukuyama and co-workers. Here we make use of the RG calculations of the vertex functions that were presented in Secs. V and VI.

We first notice that in terms of the Q-matrix formulation of Sec. III we have

$$C_{n_1 n_2}(\mathbf{q}=0, m=0) = \langle {}^1_0 q^{\alpha\alpha}_{n_1,-n_1}(\mathbf{p}=0) \times {}^1_0 q^{\alpha\alpha}_{n_2,-n_2}(\mathbf{p}=0)^* \rangle_{K^{c(s)}=0} \ , \tag{7.20}$$

where the q propagator has to be taken for a system with vanishing bare Cooper channel interaction, $K^{c(s)}=0$. If we use the Gaussian approximation for this propagator, Eq. (3.122c), we recover the usual BCS result (we recall that H is a DOS, and $K^{c(s)}$ is a DOS squared times a potential),

$$T_c = \omega_D \exp[-2H/|K^{c(s)}|] \ . \tag{7.21}$$

If we go beyond the Gaussian approximation, then the disorder generates a nonzero Cooper channel coupling constant $\delta k^{c(s)}$, even though the bare $K^{c(s)}$ vanishes. Equations (5.33) and (6.4) show that at one-loop order $\delta k^{c(s)}$ is positive, i.e., repulsive, and frequency dependent. In principle this means that the inversion problem discussed below Eq. (5.40) must be solved. This has not been done in the existing literature, where the "mean-field" approximation has been defined to include ignoring the inversion problem. While this puts the results at odds with perturbation theory, one can say in defense of this approximation that (1) in 3-d systems the frequency-dependent terms in $\delta k^{c(s)}$ are subleading compared to the leading BCS logarithms, and (2) in 2-d systems, where the frequency dependence of $\delta k^{c(s)}$ is logarithmic itself, the theory is somewhat doubtful anyway, and there are other problems that have not been addressed, which we shall come back to. Ignoring the inversion problem, the only frequency dependence is the explicit one in the propagators, which upon summation gives rise to factors of $\ln(\omega_D/T)$. In addition to the renormalization of the coupling constants, there is also a field renormalization function $\mathbb{Z}$ multiplying the propagator; see Eq. (6.6). Then the T_c equation reads

$$\frac{1}{|K^{c(s)}|} = \frac{\mathbb{Z}}{h} \frac{\ln(\omega_D/T_c)}{1+(\delta k^{c(s)}/2h)\ln(\omega_D/T_c)} \tag{7.22a}$$

or

$$T_c = \omega_D \exp\left[\frac{-2h}{\mathbb{Z}|K^{c(s)}| - \delta k^{c(s)}}\right] \ . \tag{7.22b}$$

A crude way to take into account that h, $\delta k^{c(s)}$, and $\mathbb{Z}$ are all scale dependent and therefore temperature or frequency dependent is to write the T_c equation in implicit form (Kirkpatrick and Belitz, 1992c),

$$1 = \int_{T_c}^{\omega_D} d\omega \, \gamma_c(\omega)/2\omega \ , \tag{7.23a}$$

where

$$\gamma_c(\omega) = [Z(\omega)|K^{c(s)}| - \delta k^{c(s)}(\omega)]/h(\omega) \ . \tag{7.23b}$$

The scale dependence of γ_c can be obtained from the flow equations derived in Secs. V and VI for various universality classes. Before we do so, let us discuss the physical content of the above result. For this purpose, the oversimplified version (7.22b) of the T_c equation is most convenient. Let us start out with the BCS result, Eq. (7.21), which we write as

$$T_c^0 = \omega_D \exp\left[-\frac{N_F}{N_F^2 v}\right] \ , \tag{7.24}$$

with $v>0$ an attractive effective potential. One effect of disorder that one would expect is a renormalization of the DOS factors in the exponential (Valles *et al.*, 1989). *A priori,* without a detailed analysis of the pair propagator, it is not clear which clean limit the factors of N_F actually represent. Equation (7.22b) shows that the factor N_F^2 in the denominator represents a single-particle DOS squared, since $\mathbb{Z}^{1/2}$ renormalizes that quantity [see, for example, Eq. (6.6)]. On the other hand, the factor of N_F in the numerator represents the quantity H, which has been interpreted as a quasiparticle DOS in a generalized Fermi-liquid picture (Castellani, Kotliar, and Lee, 1987); see Sec. III.B.3.d. In addition to these two different DOS effects there is $\delta k^{c(s)}$, which is a repulsive interaction in the Cooper channel and can be interpreted as a disorder-induced enhancement of the Coulomb pseudopotential (the clean-limit Coulomb pseudopotential is included in v).

At this point we can also compare the result, Eqs. (7.22) and (7.23), with the theories discussed in the preceding subsection. The work of Fukuyama *et al.* treated all three effects mentioned above in perturbation theory with respect to both disorder and the interaction strength. Anderson *et al.* stressed the importance of retardation for an enhancement of the Coulomb pseudopotential. This can also be seen from Eq. (7.23b): $\delta k^{c(s)}$ vanishes at high frequencies and increases monotonically with decreasing ω. In the generalized McMillan formula (7.5), the quantity Y' was related to the single-particle DOS, while no term related to H appeared. This is due to an approximation made by Belitz (1987c, 1987d), which amounts to neglecting certain vertex corrections. As a result, that theory could not distinguish between H and the single-particle DOS. Finally, the enhancement of λ discussed by Keck and Schmid is obviously not contained in the present model, which considers $K^{c(s)}$ a phenomenological constant. An important point to note, however, is that the RG analysis allows one to conclude that within this model the three quantities h, $\mathbb{Z}$, and $\delta k^{c(s)}$ contain the *complete* disorder dependence of T_c.

One can now determine $\gamma^c(\omega)$ and calculate T_c, Eqs. (7.23), by using the results of Secs. V and VI. Kirkpatrick and Belitz (1992c) have reported explicit results for the case of strong spin-orbit scattering, Sec. V.A.3. With Eqs. (7.23b), (5.33b), (5.36a), (5.38a), and (5.38b) the RG flow equation for γ_c is (with $l_s\to -2/\varepsilon$ and $k_s=-h$, since we consider the long-range case)

$$b\frac{d\gamma_c}{db}=-\frac{1}{2\varepsilon}g\gamma_c-\frac{1}{4}g\left[1-\frac{\gamma_c}{2}\right],\quad d=2+\varepsilon\ , \tag{7.25}$$

and the flow equations for g and h are still given by Eqs. (5.38a) and (5.38b). It is a nontrivial result that Eq. (7.25) is identical to Eq. (5.38c) with $\gamma_c\to-\gamma_c$ and that the flow equations for g and h do not depend on whether or not a bare repulsive $K^{c(s)}$ is present. To understand this, we note that $k_{c,s}$ does not appear in Eqs. (5.36a)–(5.36d). This interaction amplitude actually does appear in the perturbation theory of Sec. V.A.3, but always in some combination that vanishes as $f_n(\mathbf{p})\to\infty$ (see the discussion at the beginning of Sec. V.A.3.a). The net result is that the flow equations for γ_c, g, and h are the same with or without a bare repulsive $K_c^{(s)}$. Note also that the $f_n(\mathbf{p})\to\infty$ limit discussed in Sec. V.A.3 cannot be taken if $K_c^{(s)}<0$, and the final results given by Eqs. (5.38) are not valid in this case.

Below Eq. (7.21) we have stressed that the approach described here is rather suspect in $d=2$. If one applies it to the 2-d case anyway, the $1/\varepsilon$ in the first term on the right-hand side of Eq. (7.25) gets replaced by $\ln b$ for $b\to\infty$. For arbitrary b, a careful evaluation of Eq. (5.5a) gives

$$b\frac{d\gamma_c}{db}=-\frac{f(b)}{2}g\gamma_c-\frac{1}{4}g\left[1-\frac{\gamma_c}{2}\right],\quad d=2\ , \tag{7.26a}$$

where

$$f(b)=\frac{\pi/2}{1+(x_D b)^2}+\frac{\ln(x_D b)}{1+(x_D b)^{-2}}\ . \tag{7.26b}$$

Here $x_D=\kappa_2/q$ with κ_2 the screening wave number and q_0 the initial wave-number scale, which is on the order of the Debye wave number. The RG length scale b is related to the frequency (or temperature) by (see Sec. IV)

$$\omega/\omega_0=H/b^d h(b)\ , \tag{7.27}$$

with the flow equation for h given by Eq. (5.38b). $\omega_0=\omega(b=1)$ is the initial frequency scale, which is on the order of ω_D. Figures 44 and 45 show results for $d=3$ and $d=2$, respectively, in comparison with experimental data. These results will be discussed in Sec. VII.C. Simi-

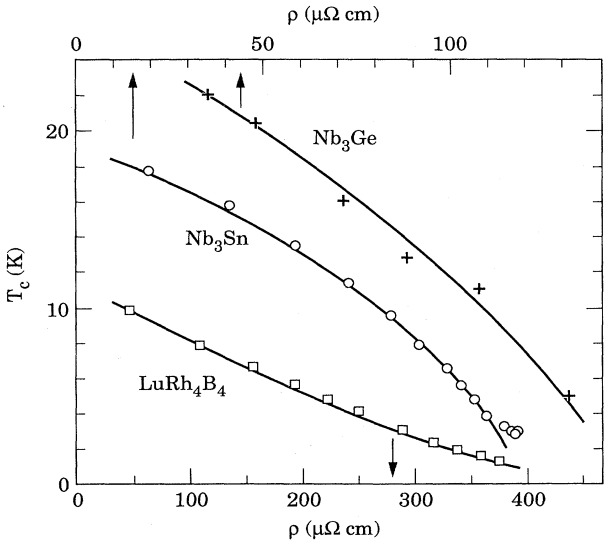

FIG. 44. Fit of Eqs. (7.23), (7.25), and (7.27) (solid curves) to experimental results of Rowell and Dynes (1980): +, on Nb_3Ge; ○, on Nb_3Sn; □, on $LuRh_4B_4$. From Kirkpatrick and Belitz (1992c).

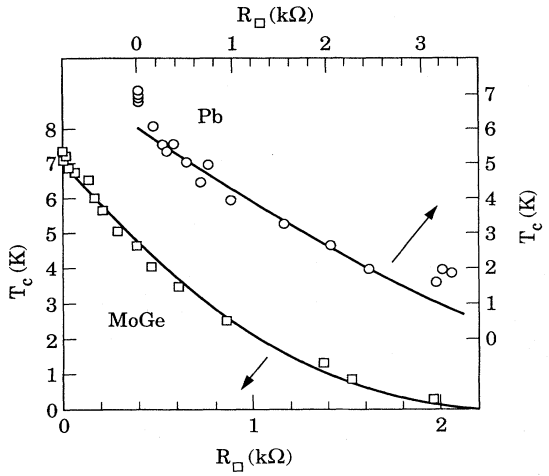

FIG. 45. Fit of Eqs. (7.23), (7.26), and (7.27) (solid curves) to experimental results on films of MoGe (Graybeal and Beasley, 1984; Graybeal, 1985) and Pb (Valles *et al.*, 1989). From Kirkpatrick and Belitz (1992c).

lar results can be obtained for the generic universality class G using Eq. (7.23b) and the results of Sec. VI.

b. The nonlinear-sigma-model approach

Finkel'stein (1987) has used the nonlinear sigma model derived in Sec. III to compute the T_c degradation due to disorder in 2-d systems with strong spin-orbit scattering. This theory has some significant advantages for describing experiments on thin-film superconductors over the theory discussed in the previous subsection. However, there still are two fundamental problems with the description of thin-film superconductors which this theory does not address either. First, in 2-d systems the superconductor-to-normal metal transition is of Kosterlitz-Thouless type. Inherent in Finkel'stein's approach, as well as in that of Kirkpatrick and Belitz, is the assumption that the transition is BCS like. Second, in the expansion about mean-field theory one encounters the fluctuation propagator,

$$\mathcal{I}_{n_1 n_2}(\mathbf{q},m)=[(\gamma^c)^{-1}(m)-\pi TC(\mathbf{q},m)]^{-1}_{n_1 n_2} . \quad (7.28)$$

By definition of the mean-field transition temperature T_c, $\mathcal{I}$ turns into a massless diffusion propagator (at small q and Ω_m) at $T=T_c$. In $d\leq 2$ this singularity leads to divergent contributions to the T_c degradation in first order in the disorder (Schmid, 1970; Strongin *et al.*, 1970; Ovchinnikov, 1973; Eckern and Pelzer, 1988). Singular contributions to T_c in $d\leq 2$ are actually expected from general phase-transition theory (see, for example, Le Bellac, 1991). In the absence of a solution of this problem, any T_c-degradation calculation is ill defined.

While Finkel'stein's approach does not address these difficulties, it has an advantage over the one discussed in the last subsection because it addresses the inversion problem mentioned below Eq. (7.21), which was just ignored in deriving, say, Eq. (7.23). Finkel'stein's basic idea is as follows. At the superconductor-to-normal metal phase transition the Cooper propagator Γ_c that was defined in Sec. V.3.b should be singular at zero momentum and $T=T_c$. Assuming that it is meaningful to derive a RG flow equation for $\tilde{\Gamma}_c=-\Gamma_c$ (remember $K^{c(s)}<0$), Finkel'stein has used the approximations described below Eq. (5.42a) to obtain

$$b\frac{d\tilde{\Gamma}_c}{db}=\tilde{\Gamma}_c^2-g/4 . \quad (7.29)$$

The transition temperature is then given by the condition $\tilde{\Gamma}_c(b_c)=\infty$, with Eq. (7.27) relating T_c and b_c. For 2-d systems it is possible to argue that it takes little disorder to suppress T_c to zero and that g in Eq. (7.29) can be regarded as a constant. Finkel'stein (1987) compared his results with the T_c data on MoGe films obtained by Graybeal and Beasley (1984). The fit is of the same quality as that shown in Fig. 45.

An interesting question is the relationship between Finkel'stein's theory and the mean-field theory discussed in the previous subsection. The latter suggests the existence of three distinct physical mechanisms that lead to T_c degradation: suppression of the single-particle DOS, change in the quasiparticle DOS, and increase of the Coulomb pseudopotential. An examination of Eq. (7.29) and the details of its derivation shows that Finkel'stein's theory does not contain either DOS effect. While Finkel'stein's approximations are consistent with simply neglecting the quasiparticle-DOS effects, the single-particle DOS effects are a more complicated matter.

In a leading-log expansion of the T_c degradation in $d=2$, the terms that lead to the log-squared anomaly in the DOS lead to a log-cubed contribution to the T_c degradation (Maekawa *et al.*, 1984). Terms of this order are also generated by the renormalization of the Coulomb pseudopotential and by contributions proportional to the fluctuation propagator. If we ignore the problems mentioned below Eqs. (7.28), the total leading perturbative result can be written in the form

$$\ln(T_c/T_c^0)=-\hat{R}_\square A[\ln(T_0/T_c^0)]^3+O\{[\ln(T_0/T_c^0)]^2\} . \quad (7.30)$$

Here $\hat{R}_\square=R_\square e^2/12\pi^2\hbar$ is the dimensionless sheet resistance, and T_0 is a microscopic temperature scale. For systems with spin-orbit scattering, the first two contributions to the prefactor A mentioned above are $A^{(1)}=2$ (Coulomb pseudopotential) and $A^{(2)}=1$ (DOS; Maekawa and Fukuyama, 1982; Maekawa *et al.*, 1984; Finkel'stein, 1988). The last contribution has not been reported separately. The calculation yields $A^{(3)}=-2$. The net result is thus $A=1$, and this is reproduced by Eq. (7.29). It is not clear what, if any, general conclusions can be drawn from the fact that there is a partial cancellation between the various contributions to A. Finkel'stein has claimed

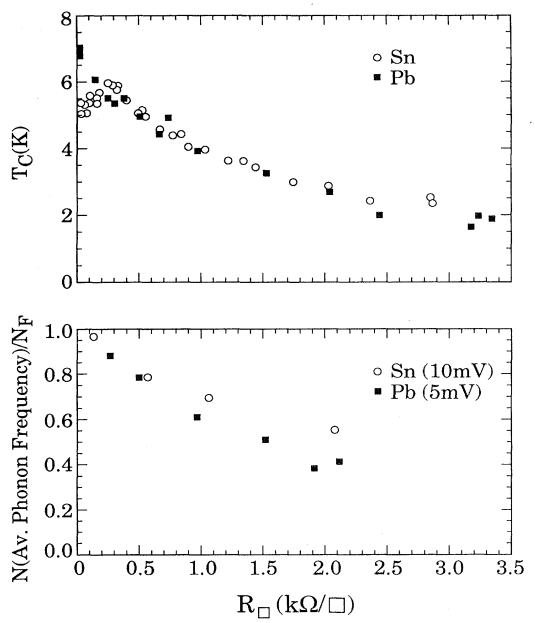

FIG. 46. T_c and tunneling DOS for Sn and Pb films as functions of sheet resistance $R_\square$. After Valles *et al.* (1989).

that the DOS effects are canceled by other terms and do not contribute to the T_c degradation. However, the issue is still controversial. Since in 2-d perturbation theory the various terms all have the same form, one could as well argue, on a perturbative level, that the net result $A=1$ reflects only the DOS term. There are also DOS contributions that are subleading in $d=2$ (i.e., they do not contribute to the leading $\ln^3$ term) and that do not get canceled by any other terms. A common misinterpretation of the published results is that there is a total cancellation of all DOS contributions to the T_c degradation. This is not correct. Finally, one has to keep in mind that terms that are equal and opposite in perturbation theory can behave very differently under renormalization.

Given the qualitative and mutually contradicting theoretical results, detailed comparisons with experiment have to be viewed with reservations. Moreover, since both approaches achieve equally good fits to experimental results, one cannot simply settle the controversy about the presence of the DOS terms by reference to experiment. Still, there is at least qualitative experimental support for DOS effects to play a role in the T_c-degradation problem, as a clear correlation has been observed between the suppression of the single-particle DOS and that of T_c in Pb and Sn films (Valles *et al.*, 1990; see Fig. 46).

c. Other approaches

A completely different approach has been pursued by Oppermann (1985, 1986, 1987a, 1987b, 1987c, 1988a, 1988b, 1988c; a summary has been given by Oppermann, 1990), who neglects all interactions except for the attractive ones in the Cooper channel. In the absence of fluctuations one can then derive a nonlinear sigma model which describes localization of quasiparticles without affecting the superconducting order (Oppermann, 1987b). This is in agreement with the observation by Ma and Lee (1985). However, Oppermann (1986) realized that situations are possible in which the average order parameter vanishes, $\langle\Delta\rangle=0$, while the order-parameter fluctuations are nonzero, $\langle\Delta^2\rangle\neq 0$, and that this gives rise to gapless superconductivity. He calls such phases superconducting glass phases. He made a symmetry analysis of models displaying such phases (Oppermann, 1987a) and found that, depending on how the symmetry of the underlying field theory is broken, one can obtain superconducting glasses that are analogous to Ising or x-y spin glasses (Oppermann, 1987c, 1988a). The density of states and the specific heat of the superconducting glass phases have also been worked out (Oppermann, 1988b, 1988c). A different and more phenomenological discussion of superconducting glass phases has been given by Fisher, Fisher, and Huse (1991).

Ramakrishnan (1989) and Kravtsov and Oppermann (1991) have argued that fluctuations of the order-parameter phase can be important even in inhomogeneously disordered systems. It is currently not known how the disorder enhancement of the repulsive interaction discussed in the last subsection will modify these effects and vice versa.

C. Experiments

As we mentioned earlier, there is a large body of experimental literature on disordered superconductors, but there are not many systematic studies of the region where superconductivity disappears. In this subsection we discuss experiments which clearly show T_c going to zero as a function of disorder. As in Sec. VII.B above, we exclude from our discussion granular systems and studies whose main objective was to investigate the Kosterlitz-Thouless or related transitions in thin films.

1. Bulk systems

Figure 44 shows theoretical results of the previous subsection together with experimental data on Nb_3Ge, Nb_3Sn, and $LuRh_4B_4$. The relation between the RG parameter $g_0=g(b=1)$ and the resistivity ρ has been taken to be $\rho=\rho^* g_0/g_c$, where $g_c=8$, and ρ^* is a resistivity scale on the order of the Mott number. The parameters are then $\gamma_c^0=\gamma_c(b=1)$, $T_c^0=T_c(\rho=0)$, and ρ^*. Good agreement is obtained with reasonable parameters. Note that the curvature of the curves changes sign as a function of γ_c^0 and that this curvature change is necessary to explain the experimental results.

We next discuss some experiments on bulk homogeneous systems that the theory of Sec. VII.B.2 *cannot* explain. Nishida *et al.* (1984; Furubayashi *et al.*, 1985),

after having found superconductivity in amorphous $Si_{1-x}Au_x$ prepared by electron-beam evaporation, studied both T_c and the extrapolated residual conductivity σ_0 as a function of Au concentration x. The material shows superconductivity in the range $0.18 \leq x \leq 0.40$ and a metal-insulator transition at $x \approx 0.14$. In the region of small σ_0 and small T_c, both quantities show a linear dependence on x within the experimental uncertainties; see Fig. 47. Similar observations have been made on amorphous Nb_xSi_{1-x} (Bishop *et al.*, 1985). The metal-insulator transition is at $x \approx 0.12$, and T_c vanishes close to $x = 0.18$ (Fig. 48). Despite the fundamental difference that Nb is a superconductor while Au is not, the virtually identical behavior of these two systems close to the metal-insulator transition makes it seem likely that the basic physics in this region is the same for both materials. In both systems superconductivity disappears only very close to the metal-insulator transition, with σ_0 on the order of 100 $(\Omega\,\mathrm{cm})^{-1}$. This is in contrast to the behavior of the A-15 and rare-earth rhodium boride materials mentioned earlier, for which an extrapolation of T_c reaches zero at substantially higher values of σ_0; see Fig. 44. It is also inconsistent with the theory of Sec. VII.B.2, which predicts that T_c will vanish far from the metal-insulator transition. The example of SiAu makes it clear that this discrepancy must be due to a very effective T_c-increasing mechanism which is not understood. Jisrawi *et al.* (1987), who have prepared superconducting AuSi samples by an ion implantation technique, have suggested that the amorphous AuSi may be similar to the superconducting high-pressure phases of Si, where superconductivity is believed to be due to a soft phonon mode (Chang *et al.*, 1985).

Similar results have been obtained in Al_xGe_{1-x} (Lesueur *et al.*, 1988) and in Mo_xGe_{1-x} (Carter, 1983). Again, in both systems T_c vanishes very close to the metal-insulator transition, and T_c varies linearly with σ_0 within the experimental error bars. Lesueur *et al.* (1988) have discussed evidence that electron-electron interactions together with disorder are important for the destruction of superconductivity. Specifically, they have observed a correlation between the Coulomb gap in the single-particle DOS and T_c. While the theory discussed in Sec. VII.B.2 above qualitatively accounts for such a correlation [see Eqs. (7.22) and the following discussion], it can explain neither why superconductivity persists so close to the metal-insulator transition, nor the linear dependence of T_c on σ_0. The reason is presumably that, as in the Si-based materials, there is a strong T_c-enhancing mechanism, whose description requires a microscopic theory of the pairing mechanism. There has been some debate about whether the pairing, at least in AlGe, is perhaps not due to electron-phonon coupling (Bernas and Nédellec, 1981).

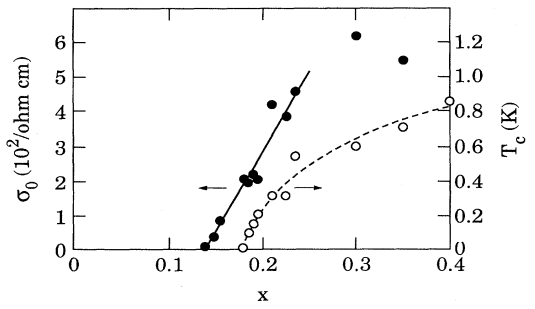

FIG. 47. Conductivity and superconducting T_c of Au_xSi_{1-x} as functions of Au concentration. The lines are guides to the eye. From Furubayashi *et al.* (1985).

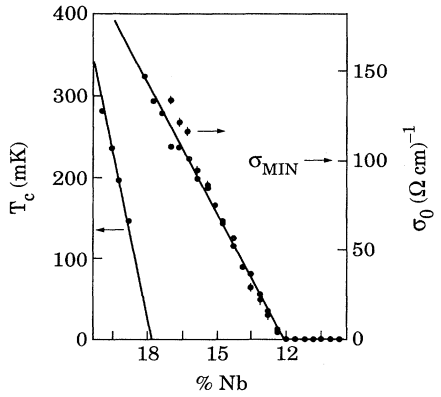

FIG. 48. Conductivity and superconducting T_c of Nb_xSi_{1-x} as functions of Nb concentration. The lines are guides to the eye. From Bishop *et al.* (1985).

2. Thin films

Preparing ultrathin, homogeneous superconducting films is difficult. This task was first accomplished by Strongin *et al.* (1970), who found that with underlayers of Ge, SiO, or Al_2O_3 they could obtain continuous films of Pb, Sn, and Bi, which were as thin as 10 Å. With decreasing film thickness, or increasing sheet resistance, they found a monotonically decreasing T_c (see Fig. 49). The actual experiment is done by starting with the thinnest film, measuring its properties *in situ*, and then depositing more and more material. The same technique has been used by Dynes *et al.* (1986) to produce Pb films as thin as 3 Å with a sheet resistance of 3.5 kΩ. Valles *et al.* (1989) have also prepared Sn films in this way and have found results for T_c that are very similar to those on Pb. The Pb data are shown in Fig. 45 together with theoretical results by Kirkpatrick and Belitz (1992c). A nominally identical system has been investigated by Haviland *et al.* (1989). The relation between $R_\square$ and nominal film thickness is different in the two experiments, and so is the relation between T_c and $R_\square$, with the sample of Haviland *et al.* showing a higher T_c at a given $R_\square$. In addition, the sample of Haviland *et al.* shows much more curvature in T_c vs $R_\square$ at $R_\square \leq 2$ kΩ

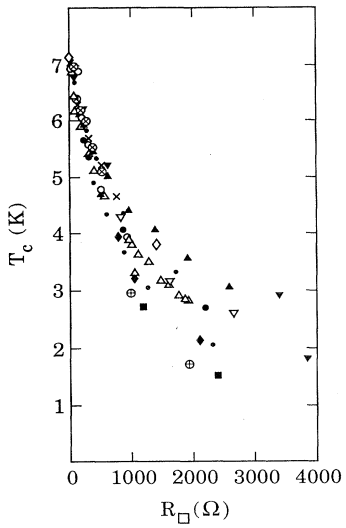

FIG. 49. T_c as a function of resistance for Pb films. ⊕, ■, Pb on Ge; +, Pb on Al_2O_3; ▽, ⊗, ○, ●, ●, △, ▲, ▼, Pb on SiO_2; all deposited at helium temperature; ♦, ◇, Pb on Ge. After Strongin *et al.* (1970).

than the experiment of Valles *et al.* The latter group observed a pronounced kink in the curve at $R_\square \approx 250\ \Omega$, which they attribute to a structural phase transition. For higher $R_\square$, both Pb and Sn show little curvature. It is not known exactly what the reason for these differences is. However, it is interesting to note that the scatter in T_c in the data of Dynes *et al.* and Valles *et al.* is due to differential sample runs' being plotted in the same figure (see also Fig. 49). For each separate run a smooth curve was obtained (Valles, 1992). Since in granular films T_c hardly decreases at all (see Fig. 43), one can speculate that the materials properties vary slightly from sample to sample, and more drastically so with overall deposition conditions. This could produce samples with varying degrees of homogeneity, with those showing the lowest T_c for a given $R_\square$ being the most homogeneous. Unfortunately, no direct efforts seem to have been made to determine the degree to which the films are homogeneous.

Homogeneous films of amorphous MoGe have been studied by Graybeal *et al.* (1984), Graybeal and Beasley (1984), and Graybeal (1985). Their thinnest superconducting sample had a nominal film thickness of 12 Å and a $R_\square$ of 1.96 kΩ. Three thinner samples with higher $R_\square$ did not show superconductivity down to 30 mK. The data from Graybeal (1985) are shown in Fig. 45 together with theoretical results. It is somewhat surprising to see the good agreement between this experiment and the theory of Sec. VII.B.2, since the theory cannot explain the experiments on bulk Ge-based materials discussed above. Results that are very similar to those on MoGe films have been obtained by Lee and Ketterson (1990) on MoC films.

VIII. DISORDER-INDUCED UNCONVENTIONAL SUPERCONDUCTIVITY

In Sec. VI we saw that the renormalization group generates an interaction $K^{c(t)}$ in the particle-particle spin-triplet channel, even if there is none in the bare theory. It was noted by Kirkpatrick and Belitz (1991) that this can lead to an unusual form of superconductivity. Some characteristic features of this kind of superconductivity have been calculated by Belitz and Kirkpatrick (1992). In this section we briefly discuss this suggestion in the general context of the preceding exposition. Details can be found in the two papers cited above.

A. A mechanism for even-parity spin-triplet superconductivity

Let us consider the particle-particle spin-triplet interaction amplitude we found at one-loop order, Eq. (6.5), and also the corresponding renormalization of the particle-particle spin-singlet amplitude, Eqs. (6.4). It is useful to rewrite these results in the following form:

$$K^{c(t),(1\text{-loop})}_{n_1 n_2, n_3 n_4} = -\delta K_t^{c(t)} - \delta K_s^{c(t)} , \tag{8.1a}$$

$$K^{c(s),(1\text{-loop})}_{n_1 n_2, n_3 n_4} - K^{c(s)} = \frac{\bar{G}}{8} \int_{\mathbf{p}} [3K^{(t)} \mathcal{D}^t_{n_3 - n_2}(\mathbf{p}) - K^{(s)} \mathcal{D}^s_{n_3 - n_2}(\mathbf{p})] + 3\delta K_t^{c(t)} - \delta K_s^{c(t)} , \tag{8.1b}$$

where

$$\delta K_t^{c(t)} = -\frac{\bar{G}}{8} \int_{\mathbf{p}} [K^{(t)} \Delta \mathcal{D}^t_{n_3 - n_2}(\mathbf{p}) + G |\Omega_{n_2} - \Omega_{n_4}| (K^{(t)})^2 \mathcal{D}^t_{|n_2 - n_4|}(\mathbf{p}) \mathcal{D}_{n_3 - n_2}(\mathbf{p})] , \tag{8.1c}$$

$$\delta K_s^{c(t)} = -\frac{\bar{G}}{8} \int_{\mathbf{p}} [K^{(s)} \Delta \mathcal{D}^t_{n_3 - n_2}(\mathbf{p}) + G |\Omega_{n_2} - \Omega_{n_4}| (K^{(s)})^2 \mathcal{D}^s_{|n_2 - n_4|}(\mathbf{p}) \mathcal{D}_{n_3 - n_2}(\mathbf{p})] . \tag{8.1d}$$

In Eqs. (8.1) we have omitted the frequency indices on $\delta K_s^{c(t)}$ and $\delta K_t^{c(t)}$. We note the following features of Eqs. (8.1).

(1) $K^{c(s),(1\text{-loop})}$ has an UV-divergent, positive (i.e., repulsive) contribution, which is lacking in $K^{c(t),(1\text{-loop})}$.

(2) $K^{c(t),(1\text{-loop})}$ has UV-finite, negative (i.e., attractive; we shall see explicitly below that $\delta K_s^{c(t)}$ and $\delta K_t^{c(t)}$ are positive semidefinite) contributions from both $K^{(s)}$ and $K^{(t)}$.

(3) $K^{c(s),(1\text{-loop})}$ has the same attractive contribution

from $K^{(s)}$ as $K^{c(t),(1\text{-loop})}$, while the one from $K^{(t)}$ is repulsive.

These features can be physically understood as follows.

(a) The UV-divergent terms in the singlet channel [cf. Eq. (6.4b)] represent the well-known disorder enhancement of the bare Coulomb repulsion, i.e., the Schmid-Altshuler-Aronov mechanism discussed in Sec. II.B. Notice that these terms are linear in the interaction amplitudes $K^{(s)}$ and $K^{(t)}$ and are therefore easily obtained in diagrammatic theory. In the triplet channel there is no bare repulsion, hence there is nothing to be enhanced, and these contributions are absent. This explains point (1) above.

(b) To understand the remaining two points, let us first consider the effect of $K^{(t)}$. By means of $K^{(t)}$, spin fluctuations are created in a Fermi liquid, i.e., an electron spin polarizes its environment. This polarization is ferromagnetic. Suppose an electron has created a spin-polarization cloud and then moves away. In a clean system the polarization cloud at a given point in space will decay exponentially with time. The polarization cloud therefore essentially moves with the electron that creates it, giving rise to the Fermi-liquid renormalization of χ_s which we discussed in Sec. III.B.3.d. In a disordered system, a spin-density fluctuation δn_s will decay algebraically, like $\delta n_s(t) \sim t^{-d/2}$, because of the long-time tail effect discussed in Sec. II.A.1.d. The polarization cloud will therefore persist even after the electron that created it has diffused away. A second electron moving into the region at a later time will still see the remains of the polarization cloud. It will get attracted to it if the two electrons form a spin triplet, and it will be repelled if they form a spin singlet. We therefore expect an attractive contribution of $K^{(t)}$ to $K^{c(t),(1\text{-loop})}$ and a repulsive one to $K^{c(s),(1\text{-loop})}$, in accord with Eqs. (8.1). Note that these effects are most pronounced if spin fluctuations decay slower than number-density fluctuations, because a time-scale separation then exists between spin and mass transport. As discussed in Sec. VI, this situation is realized in the generic universality class.

(c) Now consider $K^{(s)}$, by means of which an electron charge polarizes its environment. Again, while in a clean system this leads only to the appearance of the Fermi-liquid parameter F_0^s, in a disordered system the resulting (positive) charge-polarization cloud will decay like $\delta n_c(t) \sim t^{-d/2}$. A second electron will get attracted to this region regardless of its spin. This explains the attractive contribution of $K^{(s)}$ to both $K^{c(t),(1\text{-loop})}$ and $K^{c(s),(1\text{-loop})}$ in Eqs. (8.1).

Equations (8.1) and their interpretation above show that the net interaction in the particle-particle spin-triplet channel in a disordered system will necessarily be attractive. This raises the prospect of spin-triplet superconductivity. Before we discuss this possibility in greater detail, a few remarks are in order to put this idea in the appropriate context.

Triplet superconductivity (or superfluidity) is usually discussed by means of a static, odd-parity interaction. This is the case, for instance, in superfluid ^{3}He (e.g., Vollhardt and Wölfle, 1990) and in the Balian-Werthamer theory of triplet superconductivity (Balian and Werthamer, 1963). The effective triplet particle-particle interaction then has the antisymmetry property, Eq. (3.53), mandated by the Pauli principle. In the present case, the pairing potential has even parity, and therefore necessarily is frequency dependent. In particular, it must vanish at zero frequency. Inspection of $\delta K_{s,t}^{c(t)}$, Eqs. (8.1c) and (8.1d), shows that they go indeed as $n^{(d-2)/2}$, where n is some frequency label. Notice that $\Omega^{(d-2)/2}$ is just the Laplace transform of the $t^{-d/2}$ long-time tail. However, $K^{c(s),(1\text{-loop})}$ and $K^{c(t),(1\text{-loop})}$ as given by Eqs. (8.1) are not symmetric and antisymmetric, respectively, under interchange of n_1 and n_2 or n_3 and n_4. This does not constitute a violation of the Pauli principle, since Eqs. (8.1) have been derived as couplings in the theory in terms of the matrices q, Eqs. (3.117a) and (3.117b), which carry a restricted frequency range. Before we can use Eqs. (8.1) in the Q formulation of the theory, we therefore have to (anti)symmetrize them.

In the interpretation given above, disorder plays a crucial role in that it leads to the long-time tails which are essential for the pairing mechanism. Indeed, the strength of the pairing potential increases with increasing disorder, at least for small disorder in which the one-loop approximation is valid. This is again in sharp contrast to a Balian-Werthamer state, which is degraded by nonmagnetic disorder as effectively as a BCS state is by magnetic impurities.

It has been suggested by Béal-Monod *et al.* (1984; Béal-Monod, 1985) that the latter conclusion can be avoided in nearly ferromagnetic systems. These authors argued that strong paramagnons not only increase the triplet pairing potential, but also weaken the pair breaking by impurities. For systems with strong *antiferromagnetic* correlations, Abrahams *et al.* (1993; see also Balatsky and Abrahams, 1992) have proposed spin-*singlet* pairing that is odd in both frequency and wave number. Competition between magnetic and superconducting phases has also been studied in the context of a disordered Hubbard model (Zimanyi and Abrahams, 1990). As we mentioned in Sec. III.B.2, this kind of interplay between magnetism and superconductivity has so far not been considered within the field-theoretic models discussed in this review.

A pairing potential for triplet superconductivity that is odd in frequency and even in momentum was first proposed by Berezinskii (1974) in the context of ^{3}He. He noted that, due to the frequency dependence of the potential, extra powers of frequency or temperature appear in the BCS T_c equation [consider Eq. (7.23a) with $\gamma_c(\omega \to 0) \to 0$]. This leads to two effects. First, the interaction strength must exceed a threshold value in order for a superconducting instability to occur. Second, the system will in general reenter the normal state at very low temperatures, and the ground state will be normal. The same is true for the present mechanism in $d=3$.

However, due to the dimensionality dependence of the long-time tail exponent, the situation is qualitatively different in $d=2$. There the polarization clouds decay only like $1/t$, and in frequency space the potential is expected to be constant even at zero frequency or temperature. This is a manifestation of the strong fluctuations present in a 2-d system at low temperature. At the very least this means that in this regime strong triplet superconducting fluctuations will compete with the insulating tendencies discussed in Secs. V and VI, and with the spin instability discussed in Sec. VI. As was already mentioned in Sec. VI.A.2.c, the actual nature of the ground state of a 2-d disordered interacting electron system constitutes at present an open problem. However, the preasymptotic behavior can be worked out. This will be done in what follows in the mean-field approximation. We shall return to the 2-d ground-state problem in Sec. X.

B. Mean-field theory

We now derive a mean-field theory for spin-triplet superconductors by considering a saddle-point solution of the general field theory, Eqs. (3.92), with an appropriate ansatz for the order parameter. The derivation closely follows that of BCS-Gorkov theory given in Sec. III.B.2.b. We shall first derive the gap equation for a general pairing potential, and then separately discuss the charge- and spin-fluctuation-induced pairing mechanisms explained above.

1. The gap equation

In analogy to Eq. (3.105), we make an ansatz

$$ {}^{i}_{r}Q^{\beta\alpha}_{mn}(\mathbf{x})=\delta_{\alpha\beta}2\tau i[-\delta_{nm}\delta_{r0}\delta_{i0}\Sigma_m+\delta_{n,-m}\delta_{r1}\delta_{i1}\bar{\Delta}_m] , \tag{8.2}$$

with a real anomalous self-energy $\bar{\Delta}_m$. As a consequence of Eq. (3.66d), $\bar{\Delta}$ is an odd function of frequency,

$$\bar{\Delta}_{-m}=-\bar{\Delta}_m . \tag{8.3}$$

In principle one should consider all three spin components of ${}_1Q$ and derive coupled equations for them, as one does in the case of ^{3}He. However, all of ^{3}He's interesting features due to the three gap functions stem from the fact that the latter are momentum dependent. This is not the case here, and therefore a, say, Balian-Werthamer state and an Anderson-Brinkman-Morel state would be expected to differ at most in their spin susceptibility. The ansatz, Eq. (8.2), also neglects more general contributions to the normal self-energy, which are known to be of quantitative importance in ^{3}He.

Equation (8.2) in Eq. (3.93) yields

$$\bar{\Delta}_n=\frac{-1}{\pi N_F^2}\sum_{\mathbf{k}}T\sum_m K^{(4)}_{-n,n;m,-m}\bar{\Delta}_m+\frac{1}{2\pi N_F\tau}\sum_{\mathbf{k}}F_n^+(\mathbf{k}) , \tag{8.4a}$$

$$\Sigma_n=\frac{1}{2\pi N_F\tau}\sum_{\mathbf{k}}G_n(\mathbf{k}) , \tag{8.4b}$$

with F_n^+, G_n given by Eqs. (3.107), (3.101c), and (3.102c), with Δ in the latter equation replaced by $\bar{\Delta}_n$. The explicit disorder dependence can be eliminated in exactly the same way as in BCS-Gorkov theory. If one defines a gap function Δ by

$$\Delta_n=\bar{\Delta}_n-\frac{1}{2\pi N_F\tau}\sum_{\mathbf{k}}F_n^+(\mathbf{k}) , \tag{8.5}$$

one finally obtains the gap equation on the imaginary axis,

$$\Delta_n=T\sum_m\mathcal{K}^{(t)}_{nm}\frac{1}{N_F}\sum_{\mathbf{k}}\frac{\Delta_m}{\omega_m^2+\xi_{\mathbf{k}}^2+\Delta_m^2} . \tag{8.6a}$$

Here the kernel $\mathcal{K}^{(t)}$ is defined by

$$\mathcal{K}^{(t)}_{nm}=\frac{1}{\pi N_F}K^{(4)}_{-n-1,n;m,-m-1} , \tag{8.6b}$$

with $K^{(4)}$ from Eq. (3.92h).

While the imaginary-axis equations (8.6) are sufficient for many purposes, in general one also needs their analytic continuation to real frequencies. For sufficiently well behaved kernels $\mathcal{K}^{(t)}$ this is easily done,[19] and the result at zero temperature is

$$\Delta^R(\omega)=\frac{1}{i\pi}\int_0^\infty dx\,\mathcal{K}^{(t)}(\omega,x)\mathcal{F}^R(x) . \tag{8.7}$$

Here Δ^R and $\mathcal{F}^R$ are retarded functions, Eq. (2.11a) with $\mathcal{F}$ the Gorkov function,

$$\mathcal{F}(z)=\frac{1}{N_F}\sum_{\mathbf{k}}\frac{\Delta(z)}{\xi_{\mathbf{k}}^2+\Delta^2(z)-z^2} , \tag{8.8}$$

and $\Delta(z)$ the causal gap function for complex frequency z. $\mathcal{K}^{(t)}(\omega,x)$ is obtained from $\mathcal{K}^{(t)}_{nm}$ by the substitution $i\omega_n\to\omega+i0$, $i\omega_m\to x+i0$. $\mathcal{F}$ and Δ allow for spectral representations, Eq. (2.10a), and their reactive and dissipative parts, $\Delta',\mathcal{F}'$, and $\Delta'',\mathcal{F}''$, respectively, obey Kramers-Kronig relations,[20] Eqs. (2.12). Equation (8.3) implies the symmetry relations

$$\Delta(z)=-\Delta(-z),\quad \Delta''(\omega)=\Delta''(-\omega),\quad \Delta'(\omega)=-\Delta'(-\omega), \tag{8.9a}$$

[19]The kernel we are interested in, given by Eqs. (8.1), possesses branch cuts that make the analytic continuation very subtle. We obtain Eq. (8.7), but see Ref. 19 of Belitz and Kirkpatrick (1992) for technical difficulties encountered.

[20]This is true for physical kernels, and can be spoiled by approximations. For instance, in BCS theory $\mathcal{F}$ is causal, but Δ is not. Eliashberg theory restores the causality of Δ. In the present case the kernel is determined by causal functions, viz., diffusion poles, and both $\mathcal{F}$ and Δ are causal.

$$\mathcal{F}(z)=-\mathcal{F}(-z),\quad \mathcal{F}''(\omega)=\mathcal{F}''(-\omega),\quad \mathcal{F}'(\omega)=-\mathcal{F}'(-\omega). \tag{8.9b}$$

Since Δ_n and $\mathcal{F}_n$ are real, Eqs. (8.9) together with the spectral representation imply that Δ', Δ'', $\mathcal{F}'$, and $\mathcal{F}''$ are all purely imaginary.

2. The charge-fluctuation mechanism

Consider the singlet or charge-fluctuation contribution $\delta K_s^{c(t)}$ to $K^{c(t),(1\text{-loop})}$ in Eq. (8.1a). As mentioned earlier, we first have to antisymmetrize the interaction amplitude. We define

$$(\tilde{K}_s^{c(t)})_{n_1n_2,n_3n_4}=-\frac{1}{4}[(\delta K_s^{c(t)})_{n_1n_2,n_3n_4}-(\delta K_s^{c(t)})_{n_2n_1,n_3n_4}-(\delta K_s^{c(t)})_{n_1n_2,n_4n_3}+(\delta K_s^{c(t)})_{n_2n_1,n_4n_3}], \tag{8.10}$$

with $\delta K_s^{c(t)}$ from Eq. (8.1d). $\tilde{K}_s^{c(t)}$ has the proper antisymmetry property and can be used as $K^{(4)}$ in Eq. (8.4a). We next do the integrals in Eq. (8.1d) in $d=2$ for the case of a long-range Coulomb interaction, using Eqs. (3.134), and specialize the result according to Eq. (8.6b). To leading logarithmic accuracy we obtain

$$(\mathcal{H}_s^{(t)})_{nm}=\frac{\bar{G}H}{16\pi}\left\{F_s\left[\frac{GH}{\bar{\kappa}^2}|\omega_n+\omega_m|\right]-\left|\frac{\omega_n-\omega_m}{\omega_n+\omega_m}\right|F_s\left[\frac{GH|\omega_n-\omega_m|^2}{\bar{\kappa}^2|\omega_n+\omega_m|}\right]\right\}, \tag{8.11a}$$

where

$$F_s(x)=\frac{1}{1+x}[\ln x-\pi x^{1/2}] \tag{8.11b}$$

and

$$\bar{\kappa}=\kappa_2(1+F_0^s), \tag{8.11c}$$

with κ_2 the screening wave number from Eq. (2.45b). $\mathcal{H}_s^{(t)}$ is the singlet or charge-fluctuation contribution to the kernel in the gap equation. Actually, one should use both the charge and the spin-fluctuation contributions simultaneously in order to calculate observables. Due to various physical and technical points, however, the discussion becomes much more transparent if the two contributions are discussed separately.

The critical temperature for the charge-fluctuation mechanism is obtained as the solution of the linearized version of the imaginary-axis gap equation, Eq. (8.6a). As in BCS theory we work to leading logarithmic accuracy for $T_c\to 0$. We find

$$\Delta_n=\frac{-G}{32\pi}\ln\left[\frac{GH2\pi T_c}{\bar{\kappa}^2}\right]\sum_{m=0}^{\infty}\frac{\Delta_m}{2m+1}\left[1-\frac{|n-m|}{n+m+1}\right]. \tag{8.12}$$

In the limit $T_c\to 0$ one can further replace the frequency sum in Eq. (8.12) by an integral according to $2\pi T\sum_{m=0}^{\infty}=\int_0^{\infty}d\omega_m$. In this limit Δ_n is independent of frequency for positive or negative frequencies and has a discontinuity at zero frequency. The result for T_c can be written as

$$T_c=T_F\frac{(\bar{\kappa}/k_F)^2}{\hat{R}_\square}\exp\left[-\frac{4\pi^2/\ln 2}{\hat{R}_\square}\right]. \tag{8.13}$$

Here $\hat{R}_\square=R_\square/(\hbar/e^2)=R_\square/4108\Omega$ is the dimensionless resistance per square, and T_F is the Fermi temperature. In writing Eq. (8.13) we have used the fact that, for a 2-d system, $G=4\hat{R}_\square/\pi$, as well as the Einstein relation, Eq. (2.21).

The gap function at zero temperature can be obtained by solving Eq. (8.7). The technical details can be found in the papers quoted at the beginning of this section. The main results are as follows. As in Eliashberg theory, there are two frequency scales. The upper one is inherent to the pairing mechanism, analogous to the Debye frequency in Eliashberg theory. In the present case one expects it to depend both on the screening length and on disorder. Indeed it is given by

$$\omega_1^*=D\bar{\kappa}^2, \tag{8.14a}$$

as can be seen from Eqs. (8.11a) and (3.128c). The lower frequency scale separates high and low frequencies in the gap function and can be defined from the imaginary-axis solution by

$$\omega_2^*=\Delta(i\omega_2^*). \tag{8.14b}$$

In the limit $T_c\to 0$, the explicit result is

$$\omega_2^*=\omega_1^* e^{-1/\lambda^{1/2}}, \tag{8.14c}$$

with a coupling constant $\lambda=G/4\pi$. The analogous quantity in Eliashberg theory is the gap $\Delta=\Delta(\omega=0)$. The gap function can be obtained analytically in the limit of asymptotically small frequencies,

$$\Delta(\omega\to 0)=-i\lambda\hat{\omega}(\ln\hat{\omega})^2-\lambda\hat{\omega}(\pi+3i)\ln\hat{\omega}+O(\hat{\omega}), \tag{8.15a}$$

and in the limit of large frequencies,

$$\Delta(\omega\to\infty)=\text{const}\times\frac{\lambda}{\hat{\omega}^2}\left[\ln\hat{\omega}-2-\frac{i\pi}{2}\right]+O(\hat{\omega}^{-5/2}), \tag{8.15b}$$

where $\hat{\omega}=\omega/\omega_1^*$. The complete solution can be found numerically and is shown in Fig. 50 for $\lambda=0.1$. For coupling constants λ larger than a critical value $\lambda_c^+\simeq 0.2$ it has been shown that Eq. (8.7) does not possess a causal

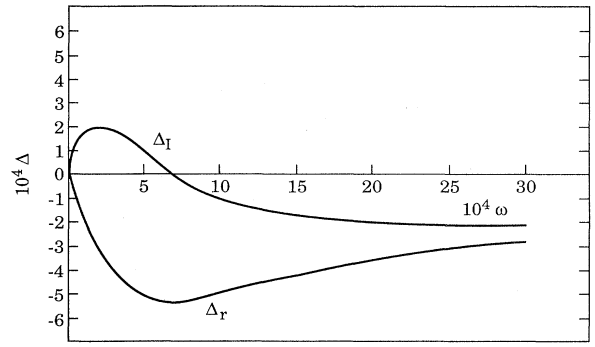

FIG. 50. Real and imaginary parts of the gap function Δ vs frequency ω for the charge-fluctuation mechanism. Δ and ω are measured in units of ω_1^*, and the coupling constant has been chosen as $\lambda=0.1$. From Belitz and Kirkpatrick (1992).

solution. The reason for, and the physical implication of, this absence of a solution is not clear at present. It is possible that the effect is an artifact due to the omission of the feedback of the superconductivity on the pairing potential. This feedback or self-consistency effect should be taken into account whenever one deals with a purely electronic mechanism for superconductivity, as has been pointed out by Pao and Bickers (1992).

The gap function determines the tunneling DOS, $N(\omega)$, through

$$\frac{N(\omega)}{N_F}=\mathrm{Re}\left[\frac{\omega}{[\omega^2-\Delta^2(\omega)]^{1/2}}\right] . \tag{8.16}$$

The asymptotic behavior is easily obtained from Eqs. (8.15), and the numerical solution is shown in Fig. 51. Finally, $N(\omega)$ determines the low-temperature specific heat via

$$C(T)=\frac{1}{T^2}\int_0^\infty d\omega\, N(\omega)\omega^2 e^{-\omega/T} , \tag{8.17}$$

which is also shown in Fig. 51.

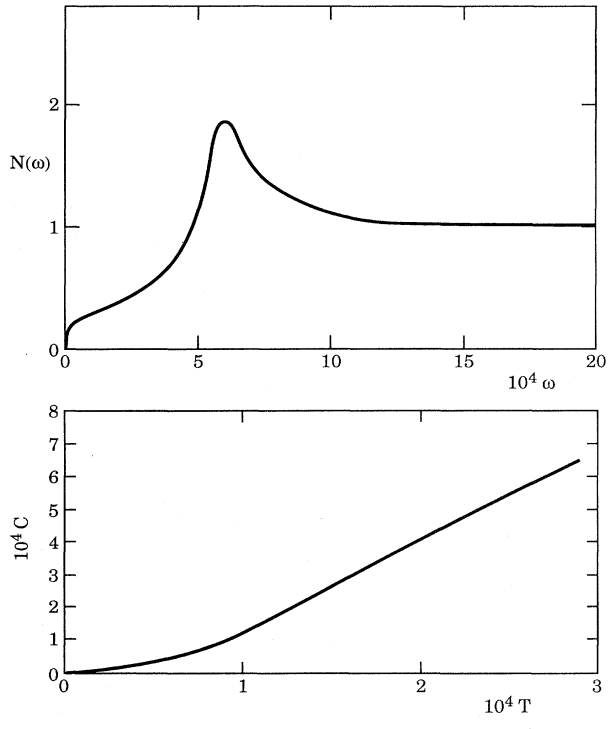

FIG. 51. Tunneling density of states N vs frequency ω, and specific heat C vs temperature T, for the gap function shown in Fig. 50. N is measured in units of N_F, C in units of $N_F\omega_1^*$, and ω and T in units of ω_1^*. From Belitz and Kirkpatrick (1992).

3. The spin-fluctuation mechanism

Now we turn to the spin-fluctuation contribution $\delta K_t^{c(t)}$ to $K^{c(t),(1\text{-loop})}$ in Eq. (8.1a). It has to be antisymmetrized according to Eq. (8.10), and performing the integrals yields

$$(\mathcal{H}_t^{(t)})_{nm}=\frac{-\bar{G}K^{(t)}}{16\pi}F_t(|\omega_n+\omega_m|/|\omega_n-\omega_m|) , \tag{8.18a}$$

where

$$F_t(x)=\frac{1}{x(1+\gamma_t^0)-1}[(x-1)\ln(1+\gamma_t^0)-\gamma_t^0 x\ln x] , \tag{8.18b}$$

with $\gamma_t^0=K^{(t)}/H$. The appearance of the triplet interaction amplitude γ_t^0 in the kernel raises both problems and interesting prospects having to do with renormalization. In general one would expect the renormalized coupling constants to enter the kernel. In the case of γ_t we know from Sec. VI that it grows under renormalization, although the precise behavior at large scales in $d=2$ is not known. We therefore replace γ_t^0 in Eq. (8.18b) by γ_t and consider the limit $\gamma_t\to\infty$. The linearized imaginary-axis gap equation, Eqs. (8.6), then reads

$$\Delta_n=\frac{y_0}{16}\sum_{m=0}^{\infty}\frac{\Delta_m}{2m+1}\ln\left|\frac{n+m+1}{n-m}\right| , \tag{8.19a}$$

with

$$y_0=\bar{G}K^{(t)}/H=\frac{2}{\pi^2}\gamma_t^0\hat{R}_\square . \tag{8.19b}$$

In the same spirit, y_0 in Eqs. (8.19) should be replaced by its scale-dependent counterpart, $y(T)$. $y(T)$ is determined by the flow equations (6.52) or (6.53a), depending on the value of γ_t^0. For sufficiently large values of γ_t^0, Eq. (6.53a) is appropriate, and one finds for the critical temperature

$$T_c=T_0\exp\left[\frac{-2}{y_0}(1-y_0/y_c)\right] , \tag{8.20}$$

where T_0 is the initial, microscopic temperature scale, which one expects to be on the order of the Fermi tem-

perature, and $y_c = 64/\pi^2$. Comparison of Eqs. (8.20) and (8.13) shows that the spin-fluctuation mechanism allows, in principle, for substantially higher values of T_c than the charge-fluctuation mechanism. We shall come back to this in the following subsection.

The gap function at zero temperature, the tunneling DOS, and the low-temperature specific heat for the spin-fluctuation mechanism have also been calculated in exact analogy to Sec. VIII.B.2 above. The results were found to be very similar to those for the charge-fluctuation mechanism shown in Figs. 50 and 51. Only the scales and some details of the asymptotic behavior are different.

C. Is this mechanism realized in nature?

Even-parity spin-triplet superconductivity has so far not been experimentally observed. On the other hand, the mechanism discussed above is so general that one would expect it to be realized in some systems. Let us briefly discuss the most uncertain points in the prediction, in which systems one might want to look for this kind of superconductivity, and what values of T_c one might expect.

Of the two mechanisms discussed, the spin-fluctuation mechanism is expected to lead to a much higher T_c than the charge-fluctuation mechanism. The physical reason for this was explained in point (b) below Eq. (8.1d). Technically, the higher T_c is due to the fact that y or γ_t is expected to increase as the temperature decreases. Taking this effect into account has led to an uncertainty in our theoretical discussion of T_c. We note that, in principle, G and H occurring in the charge-fluctuation mechanism should also be renormalized, but their scale dependence is much weaker than that of γ_t; see the discussion in Sec. VI. We conclude that a relatively high T_c should be possible via the spin-fluctuation mechanism if one can find a 2-d system that is nearly ferromagnetic, i.e., with a large value of γ_t^0, with no or very weak spin-orbit scattering. Such a material may be hard to find. If it can be found, one expects T_c to be a very rapidly varying function of both disorder and γ_t^0 (or the Fermi-liquid parameter F_0^a), and superconductivity to be observable only in a narrow region of parameter space.

For the charge-fluctuation mechanism, reasonably reliable predictions appear to be somewhat easier to make. Here the main uncertainties are twofold. First, higher-order renormalization effects have been neglected. Since the particle-particle triplet coupling constant appears only at one-loop order, one should go back to the original action, add this interaction, and repeat the renormalization process. This has not been done so far, i.e., the triplet pairing mechanism has been discussed only at the Gaussian or tree level. From our experience with the degradation effects for conventional superconductivity discussed in Sec. VII, we expect T_c degradation effects here as well, starting at two-loop order. This will change the behavior of T_c as a function of $\hat{R}_\Box$ and make T_c for large $\hat{R}_\Box$ fall off much faster than Eq. (8.13) predicts. Secondly, normal self-energy effects have been neglected so far. The latter are known to be of great quantitative importance in the case of superfluidity in ^{3}He (Levin and Valls, 1978). In fact, the omission of these effects was one of the reasons why early estimates of T_c in ^{3}He were much too high. With all these caveats in mind, one can try to estimate T_c values from Eq. (8.13) if one is careful not to allow too large a value for $\hat{R}_\Box$. A suitable system seems to be the inversion layer in a Si MOSFET (Ando *et al.*, 1982), because it is truly two dimensional, the relevant parameters are well known, and $R_\Box$ can be changed continuously by simply changing the gate voltage. Belitz and Kirkpatrick (1992) have estimated T_c for this system to lie in the range 15–640 μK.

IX. OTHER THEORETICAL APPROACHES AND RELATED TOPICS

In this section we first discuss some approaches to the interacting disordered electron problem that do not make use of the generalized nonlinear sigma model derived in Sec. III. We do not intend to review these various theories thoroughly; rather, we only sketch the main technical aspects of each method and then give the main results. We also discuss possible connections between these theories and the sigma-model approach used in the rest of this review. We then discuss two delocalization transitions that occur in 2-d disordered electronic systems, viz., the transition from an insulator to a superconductor in a thin film and the delocalization transition in the integer quantum Hall effect.

A. Results in one dimension

In the absence of interactions all states are localized in 1-d disordered electronic systems (Mott and Twose, 1961; Borland, 1963). Typical localization lengths are on the order of the elastic mean free path between electron impurity collisions. The existence of interactions between the electrons can change this situation qualitatively. In particular, for sufficiently strong attractive interactions there is a competition between superconducting fluctuations and disorder, which can lead to delocalization (Chui and Bray, 1977; Apel, 1982; Apel and Rice, 1982; Suzumura and Fukuyama, 1983, 1984). The resulting phase transition is similar to the zero-temperature superconductor-insulator transition that is expected to take place in 2-d electronic systems (Fisher, 1990; Hebard and Paalanen, 1990). The main distinction is believed to be that the 1-d zero-temperature superconducting state has only quasi-long-ranged order (i.e., algebraically decaying correlations), unlike 2-d systems, which can have true long-ranged order at $T=0$. This distinction is due to the fact that the lower critical dimension for a zero-temperature superconductor-insulator transition is one space (plus one time) dimension.

To describe this phase transition in $d=1$, a perturbative RG description specially designed for 1-d systems has been developed (Giamarchi and Schulz, 1988). A motivation for this work was the expectation that the behavior in $d=1$ would shed some light on the, presumably technically harder, problem of disordered interacting electrons in higher dimensions. The starting point is the standard "g-ology" description of 1-d interacting electron systems (Emery, 1979), which is then modified to include disorder. The treatment of the latter is modeled after Berezinskii's (1973; Abrikosov and Ryzhkin, 1978) theory of noninteracting 1-d systems. The main assumption is that the only important processes are those near the Fermi surface (which is assumed to exist). In this case the energy spectrum can be linearized around k_F, and the interaction can be parametrized by the four coupling constants $g_{1\perp}$, $g_{1\parallel}$, $g_{2\perp}$, $g_{2\parallel}$. The g's describe processes with momentum transfer close to zero ($g_{2\perp}$, $g_{2\parallel}$) and $2k_F$ ($g_{1\perp}$,$g_{1\parallel}$), respectively, and the symbols $\perp$ and $\parallel$ refer to the spin structure of the interaction. The interaction between the electrons and impurities can be parametrized by two uncorrelated Gaussian random fields η and ξ. With L denoting the system size the Hamiltonian is

$$\begin{aligned}\hat{H}=&\sum_{k,\sigma,r} v_F(rk-k_F)\hat{a}^{+}_{r,k,\sigma}\hat{a}_{r,k,\sigma}+L^{-1}\sum_{\substack{k_1,k_2,p\\ \sigma,\sigma'}}(g_{1\parallel}\delta_{\sigma,\sigma'}+g_{1\perp}\delta_{\sigma,-\sigma'})\hat{a}^{+}_{+,k_1,\sigma}\hat{a}^{+}_{-,k_2,\sigma'}\hat{a}_{+,k_2+p,\sigma'}\hat{a}_{-,k_1-p,\sigma}\\ &+L^{-1}\sum_{p,\sigma,\sigma'}(g_{2\parallel}\delta_{\sigma,\sigma'}+g_{2\perp}\delta_{\sigma,-\sigma'})\hat{\rho}_{+,\sigma}(p)\hat{\rho}_{-,\sigma'}(-p)\\ &+\sum_{\sigma}\int dx\,\eta(x)[\hat{\psi}^{+}_{+,\sigma}(x)\hat{\psi}_{+,\sigma}(x)+\hat{\psi}^{+}_{-,\sigma}(x)\hat{\psi}_{-,\sigma}(x)]\\ &+\sum_{\sigma}\int dx[\xi(x)\hat{\psi}^{+}_{+,\sigma}(x)\hat{\psi}_{-,\sigma}(x)+\xi^{*}(x)\hat{\psi}^{+}_{-,\sigma}(x)\hat{\psi}_{+,\sigma}(x)]\,, \end{aligned} \tag{9.1}$$

where

$$\hat{\rho}_{r,\sigma}(p)=\sum_{k}\hat{a}^{+}_{r,k+p,\sigma}\hat{a}_{r,k,\sigma}\,, \tag{9.2a}$$

$$\hat{\psi}_{r,\sigma}(x)=L^{-1/2}\sum_{k}\hat{a}_{r,k,\sigma}e^{ikx}\,, \tag{9.2b}$$

$r=\pm$ denotes right- and left-going fermions, $\sigma=\pm$ indicates spin up and spin down, and $\hat{a}_{r,k,\sigma}$ ($\hat{a}^{+}_{r,k,\sigma}$) is the annihilation (creation) operator for a fermion in state (r,σ) with momentum k. The fields η and ξ in Eq. (9.1) describe the forward and backward scattering of electrons by the impurities. They are governed by the Gaussian probability distributions,

$$P_{\eta}\sim\exp\left\{-D_{\eta}^{-1}\int dx\,\eta^{2}(x)\right\}\,, \tag{9.3a}$$

$$P_{\xi}\sim\exp\left\{-D_{\xi}^{-1}\int dx\,\xi^{*}(x)\xi(x)\right\}\,, \tag{9.3b}$$

where $D_{\nu}=v_F/\tau_{\nu}$, $\nu=\eta,\xi$ with τ_{ν} the scattering time associated with each process. The field η is real, whereas ξ is complex, ξ and ξ^{*} being associated with momentum transfer $2k_F$ and $-2k_F$, respectively. The effects of the forward-scattering random field η turn out not to be of qualitative importance. This feature is analogous to the irrelevance of $S^{(1)}_{\rm dis}$ in Eqs. (3.46) in the σ-model approach. The effect of the electron impurity backward scattering, as described by the random field ξ, is much more drastic. In the noninteracting limit this term leads to localization. In the interacting case both impurity backscattering and the interaction described by $g_{1\perp}$ give rise to divergent contributions in a perturbative calculation. To interpret these singularities, RG ideas can be used. To first order in the disorder, or in D_{ξ}, and to second order in $g_{1\perp}$ the RG flow equations are

$$\frac{dK_{\rho}}{d\ln b}=-\frac{1}{2}K_{\rho}^{2}\frac{u_{\rho}}{u_{\sigma}}\mathcal{D}\,, \tag{9.4a}$$

$$\frac{dK_{\sigma}}{d\ln b}=-\frac{1}{2}K_{\sigma}^{2}[\mathcal{D}+y^{2}]\,, \tag{9.4b}$$

$$\frac{du_{\rho}}{d\ln b}=-\frac{u_{\rho}^{2}K_{\rho}}{2u_{\sigma}}\mathcal{D}\,, \tag{9.4c}$$

$$\frac{du_{\sigma}}{d\ln b}=-\frac{u_{\sigma}K_{\sigma}}{2}\mathcal{D}\,, \tag{9.4d}$$

$$\frac{dy}{d\ln b}=2(1-K_{\sigma})y-\mathcal{D}\,, \tag{9.4e}$$

$$\frac{d\mathcal{D}}{d\ln b}=[3-K_{\rho}-K_{\sigma}-y]\mathcal{D}\,, \tag{9.4f}$$

with ($\nu=\rho,\sigma$),

$$K_{\nu}=\left[\frac{2\pi v_F+g_{\nu}}{2\pi v_F-g_{\nu}}\right]^{1/2}\,, \tag{9.5a}$$

$$u_{\nu}=\left[v_F^{2}-\frac{g_{\nu}^{2}}{4\pi^{2}}\right]^{1/2}\,, \tag{9.5b}$$

$$g_{\rho}=g_{1\parallel}-g_{2\parallel}-g_{2\perp}\,, \tag{9.5c}$$

$$g_{\sigma}=g_{1\parallel}-g_{2\parallel}+g_{2\perp}\,, \tag{9.5d}$$

$$\mathcal{D}=\frac{2D_{\xi}\alpha}{\pi u_{\sigma}^{2}}(u_{\sigma}/u_{\rho})^{K_{\rho}}\,, \tag{9.5e}$$

$$y = g_{1\perp}/\pi u_\sigma \ , \tag{9.5f}$$

where α is a length-scale cutoff. The labels ρ and σ denote interaction parameters related to the charge and spin degrees of freedom, respectively, and we note that in the absence of disorder ($\mathcal{D}=0$) the charge and spin degrees of freedom are not coupled.

Equations (9.4) have the following interesting features. First, the parameters K_ν and u_ν ($\nu=p,\sigma$) appear naturally in the bosonic representation of the 1-d fermion problem. In the clean limit they determine how the decoupled spin and charge correlation functions decay in space and time. Secondly, note the structure of Eqs. (9.4e) and (9.4f). $\mathcal{D}$ is essentially the disorder, and to $O(\mathcal{D})$ its renormalization is given by Eq. (9.4f). Comparing this equation with the analogous sigma-model result to $O(g)$, $dg/d\ln b = g$ (for $\varepsilon = -1$, cf. Secs. V and VI), we see a qualitative difference: the disorder is renormalized by the interactions in Eq. (9.4f) even to lowest order in $\mathcal{D}$. Similarly, we see that the renormalization of the interaction strength y given by Eq. (9.4e) contains a disorder-independent term. Both of these effects, which have no analogies in the sigma-model approach, are due to the nontrivial nature of the clean 1-d interacting ground state, where long-ranged correlations exist even in the absence of disorder or diffusion. To illustrate the complicated situation in clean 1-d systems, we show a typical phase diagram in Fig. 52 (Giamarchi and Schulz, 1988; see also Emery, 1979). The labeling in Fig. 52 indicates that the respective regions are dominated by divergent correlations of $2k_F$ charge-density-wave (CDW), spin-density-wave (SDW), singlet superconductivity (SS), and triplet superconductivity (TS) type. This phase diagram assumes vanishing spin anisotropy, in which case K_σ and u_σ can be related to y. We stress that in higher dimensions, where the sigma-model approach has been applied, it has been assumed that the clean ground state is a simple Fermi liquid with no long-ranged correlations. The validity and generality of this assumption is not clear.

Equations (9.4) have been analyzed for vanishing spin anisotropy ($g_{1\perp}=g_{1\parallel}=g_1$, $g_{2\perp}=g_{2\parallel}=g_2$). A phase diagram for finite disorder has been derived under the assumptions that runaway RG trajectories indicate flow towards strong coupling and that there are no nontrivial fixed points at intermediate coupling (Giamarchi and Schulz, 1988). The resulting phase diagram is shown in Fig. 53. In the regions labeled RAF (random antiferromagnet) and PCDW (pinned charge-density wave), $\mathcal{D}$ flows to infinity and the electrons are localized. The two possibilities are related to different magnetic properties, which are in turn determined by the sign of the fixed-point value g_1^* of g_1. If $g_1^* > 0$, then there is a repulsion between electrons of the same spin. An electron thus avoids localizing close to one in the same spin state. Due to the randomness, the exchange interaction between adjacent electrons will also be random. One therefore expects properties typical of a random antiferromagnet. If $g_1^* < 0$, then there is a nonmagnetic system of localized pairs of spins. This has been dubbed a pinned charge-density wave or a charge-density-wave glass. In the region labeled TS in Fig. 53, $\mathcal{D}$ and y flow to zero, while in the region labeled SS, $\mathcal{D}$ flows to zero and y flows to $-\infty$. In both cases the states are delocalized. Details then determine whether the dominant superconducting fluctuations are of the triplet or singlet type.

It should be noted that the phase diagram discussed above does not really follow from the flow equations given. The first-order RG flow equations imply that for the localized phase ($\mathcal{D}\to\infty$) one has $g_1\to-\infty$, i.e., a nonmagnetic behavior of the pinned charge-density-wave type, regardless of the strength of the bare g_1. This behavior arises because only first-order terms have been retained, even though both $\mathcal{D}$ and $|g_1|$ flow to infinity. In Sec. VI we discussed some of the problems that are encountered in interpreting RG equations with runaway flow trajectories. In the present case it turns out that magnetic (SDW) fluctuations interact with the disorder via second-order terms. It has been assumed that retain-

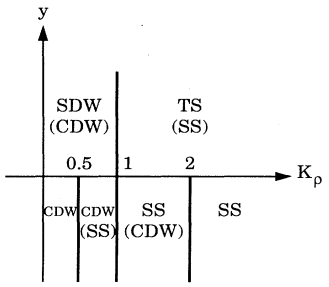

FIG. 52. Phase diagram for a clean 1-d electronic system in the $y-K_\rho$ plane. In $d=1$, the "phase diagram" means a diagram of dominant fluctuations. Less divergent fluctuations than the dominant ones are within parentheses. The respective regions are dominated by divergent correlations of $2k_F$ charge-density-wave (CDW), spin-density-wave (SDW), singlet superconductivity (SS), and triplet superconductivity (TS) type. From Giamarchi and Schulz (1988).

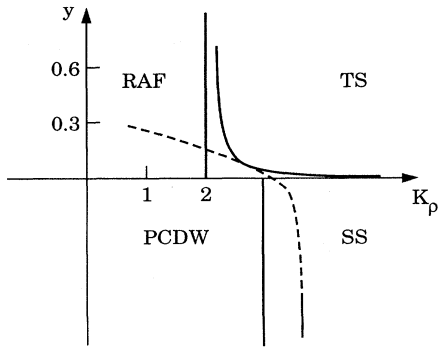

FIG. 53. Same as Fig. 52, but for finite disorder, $\mathcal{D}=0.05$. The dashed lines are parts of the phase boundaries that cannot be derived from a one-loop calculation. The two vertical lines, together with the K_ρ axis, represent the phase boundaries in the limit of vanishing disorder, $\mathcal{D}\to 0$. Note that the behavior in the limit $\mathcal{D}\to 0$ is different from that in the clean case, $\mathcal{D}=0$, shown in Fig. 52. From Giamarchi and Schulz (1988).

ing these and higher-order terms will lead to a phase diagram like that shown in Fig. 53.

The superconductor-insulator transition in Fig. 53 occurs for strongly (strength comparable to the bandwidth) attractive interactions only. For example, from Eqs. (9.5) we see that $K_\rho = 2$ means that $g_2 = -(3\pi v_F)/5$. Even for systems with a superconducting ground state such a strong attraction is unlikely to exist. In real quasi-1-d materials, the observed absence of localization is presumably due to some type of 3-d coupling between the chains.

B. Local magnetic-moment effects

Local magnetic moments are known to exist and to be important in determining materials properties in many clean electronic systems. There are two key questions in local-moment physics. The first has to do with local-moment formation. This problem was first studied by Friedel (1956) and by Anderson (1961) and Wolff (1961). The second question deals with the behavior of local moments and whether they can exist at very low temperatures. In the metallic phase this second aspect is the Kondo problem (Kondo, 1964; for an elementary discussion of the physics involved, see Anderson, 1984, pp. 188 ff.). It has been known for some time that for temperatures less than the Kondo temperature T_K a type of screening takes place by which a conduction electron forms a singlet state with a local moment making the composite object a simple nonmagnetic impurity. That is, for $T < T_K$ local moments effectively do not exist in clean materials.

In the presence of disorder two new concepts related to these problems have recently been introduced: (1) disorder can facilitate local-moment formation; (2) in some sense the Kondo temperature is zero in disordered systems. Both of these effects can have important experimental and theoretical consequences.

In this subsection we review recent work on local moments in disordered electronic systems. From this discussion it will become clear that many aspects of this problem are not incorporated in the sigma-model approach that was discussed in Secs. V and VI.

1. Local moments in the insulator phase

Many of the ideas developed to understand local moments in the metallic phase of disordered electronic systems were motivated by the behavior of local moments in the insulating phase, which was studied earlier. We begin our discussion of local moments by reviewing some of this work. To be specific we consider models of doped semiconductors (e.g., Si:P, CdS:In) at donor concentrations n below the metal-insulator transition. Because the Bohr radius of the donor electron, a, is much greater than the host lattice spacing and sets the range of the exchange coupling J, doped semiconductors can be modeled by a system of spin-$\frac{1}{2}$ objects distributed randomly in space. The magnetic properties of the system are characterized by the quantum antiferromagnetic Heisenberg Hamiltonian,

$$\hat{H} = \frac{1}{2} \sum_{i \neq j} J(\mathbf{r}_i - \mathbf{r}_j) \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j \ , \tag{9.6a}$$

where

$$J(r) = J_0 \exp(-2r/a) \ , \tag{9.6b}$$

$J_0 > 0$, falls off exponentially with r.

Building upon earlier one-dimensional work (Dasgupta and Ma, 1980), Bhatt and Lee (1982) have used numerical methods and RG ideas to study the low-temperature properties of systems described by Eq. (9.6). The solution method takes advantage of the broad distribution of exchange couplings (typically four to eight decades) to iteratively discard high-lying excitation levels of the system that are not important at low temperature and transform the Hamiltonian to a scaled version with the same low-lying states. The magnetic properties of the system can then be extracted from the distribution of energy levels. The basic physical idea behind the procedure is that, for a given T, spins connected by a $J > T$ will combine to form a nonmagnetic singlet state and drop or condense out of the problem. In some respects this is an insulator analog of Kondo screening.

At each iteration, the highest energy levels discarded are equal to J_M, the largest coupling in the system. For the highly disordered system under consideration, in which the couplings and energy levels span several orders of magnitude, it is a good approximation to consider, at $T = J_M$, all discarded levels as frozen and the remaining spins as free. This yields a magnetic susceptibility

$$\chi(T = J_M) = N(J_M) \chi_c(J_M) \ , \tag{9.7}$$

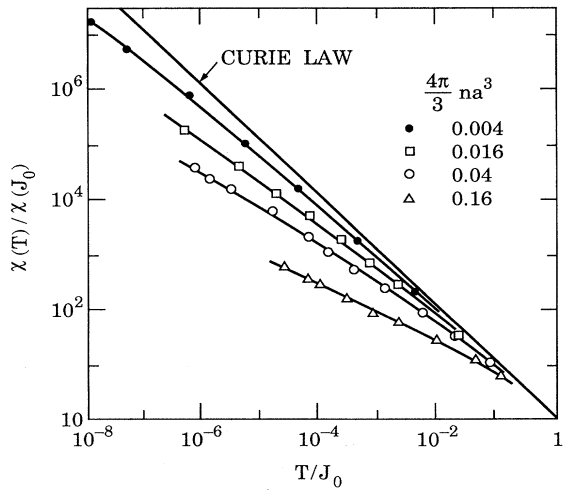

FIG. 54. Magnetic susceptibility of the spatially random three-dimensional spin-$\frac{1}{2}$ Heisenberg model with exponential exchange for various concentrations n. From Bhatt and Lee (1982).

where $\chi_c(J_M)=\mu_B^2/J_M$ is the Curie susceptibility of a free spin at temperature $T=J_M$ and $N(J_M)$ is the effective number of free spins remaining when the largest surviving coupling equals J_M.

Figure 54 shows the normalized magnetic susceptibility as a function of temperature. The Curie susceptibility is also shown. It is important to note there is no sign of any saturation as $T\to 0$. This indicates the absence of any magnetic ordering at finite T. Also note that the curves seem to suggest a sub-Curie divergence: $\chi(T\to 0)\sim T^{-a}$ with $0<a<1$. One concludes that quantum fluctuations prevent the classically expected spin-glass ordering. The zero-temperature ground state appears to be a random singlet state with a divergent magnetic susceptibility.

2. The disordered Kondo problem

In clean metallic systems with local moments, the magnetic interactions discussed in the previous subsection are thought to be less important than the interactions between the local moments and the conduction electrons, i.e., the Kondo problem. Before turning to the disordered problem and seeing if this remains true, we first recall some important features of the Kondo effect in clean systems (see, for example, Grüner and Zawadowski, 1974; Nozières, 1979). The standard Kondo Hamiltonian for a single spin-$\frac{1}{2}$ magnetic impurity at the origin is

$$\hat{H}=\sum_{\mathbf{k},\sigma}\varepsilon_{\mathbf{k}}\hat{a}^{+}_{\mathbf{k}\sigma}\hat{a}_{\mathbf{k}\sigma}+J\hat{\mathbf{S}}\cdot\hat{\mathbf{s}}(0)$$
$$=\sum_{\mathbf{k},\sigma}\varepsilon_{\mathbf{k}}\hat{a}^{+}_{\mathbf{k}\sigma}\hat{a}_{\mathbf{k}\sigma}+\frac{J}{N}\sum_{\sigma,\sigma'}\hat{\mathbf{S}}\cdot\vec{\sigma}_{\sigma\sigma'}\sum_{\mathbf{k},\mathbf{k}'}\hat{a}^{+}_{\mathbf{k}\sigma}\hat{a}_{\mathbf{k}'\sigma'}\ , \quad (9.8)$$

with $\mathbf{s}(0)$ the effective spin at the origin due to the conduction electrons, $\vec{\sigma}$ the Pauli matrices, $\hat{\mathbf{S}}$ the impurity spin, and $J>0$ an antiferromagnetic coupling. Ordinary perturbation theory and perturbative RG calculations imply that the renormalized antiferromagnetic coupling becomes very large below a "Kondo temperature" T_K given by

$$T_K=E\exp[-1/JN_F]\ , \quad (9.9)$$

where E is an energy scale on the order of the bandwidth or the Fermi energy. Our focus will be on the exponential dependence of T_K on JN_F in Eq. (9.9). The following physical picture emerges. For $T>T_K$ there are local moments that give a Curie-like contribution to the magnetic susceptibility,

$$\chi_{\mathrm{LM}}(T>T_K)\sim\frac{\mu^2}{T}n_I\ , \quad (9.10a)$$

with μ as the magnetic moment of a local moment and n_I the density of local moments. For $T<T_K$ the renormalized antiferromagnetic coupling scales to infinity, and it is energetically favorable for the impurity spin to "trap" a conduction electron. One then has a spin-singlet composite object with nonmagnetic properties. The local-moment contribution to χ is thus cut off at T_K, and for lower temperatures it is given by

$$\chi_{\mathrm{LM}}(T<T_K)\sim\frac{\mu^2}{T_K}n_I\ . \quad (9.10b)$$

The crucial physical point is that for $T<T_K$ the local moment is effectively screened, and for $T\to 0$ the system is a conventional Fermi liquid. Note that, in these arguments, the effects of local-moment/local-moment interactions are ignored. This is an important current research topic, and the effects of these interactions are not yet understood.

We next argue that a generic *disordered* system with local moments will *not* be a Fermi liquid at $T\to 0$. We then discuss some of the possible consequences of this non-Fermi-liquid behavior for the metal-insulator transition. There are two distinct mechanisms that can lead to the absence of Kondo screening in disordered systems. The first (Bhatt and Fisher, 1992) is a local phenomenon and rests on the statistical properties of the random potentials that are important in disordered electronic systems. Consider a local moment at the origin. With a probability

$$P(r)\sim e^{-cr^d}\ , \quad (9.11a)$$

the local moment is located in a cavity of radius r and isolated from the rest of the system. It is plausible to assume that the exchange coupling between the local moment and the rest of the system, with the nearest neighbor a distance r away, goes as

$$J(r)\sim e^{-c'r}\ . \quad (9.11b)$$

In Eqs. (9.11) c and c' are constants. In the absence of J the local moment would give a divergent Curie susceptibility, $\chi\sim\mu^2/T$. However, the coupling to the conduction electrons can quench or screen the local moment.

In general a local moment will not be Kondo quenched if J is too small, i.e., if r in Eq. (9.11b) is too large. A critical value for r can be determined by using Eq. (9.9) with $J\to J(r)$ and $T_K\to T_K(r)$. That is, in a disordered system there will be a range or distribution of Kondo temperatures. Using this, one concludes that an isolated local moment will not be Kondo screened at temperature T, if $r>r_c$ with

$$T=E\exp[-e^{c'r_c}/N_F]\ , \quad (9.12a)$$

or

$$r_c\sim\ln\left[\ln\frac{E}{T}\right]\ . \quad (9.12b)$$

The density of unquenched local moments, $n_I(T)$, is then proportional to $P(r_c)$ given by Eqs. (9.11a) and (9.12b):

$$n_I(r_c)=n_I(T)\sim\exp\left\{-c''\left[\ln\left[\ln\frac{E}{T}\right]\right]^d\right\}\ , \quad (9.13)$$

with a constant c''. This leads to a local-moment mag-

netic susceptibility given by

$$\chi_{\rm LM}(T)\sim\frac{\mu^2}{T}\exp\left\{-c''\left[\ln\left[\ln\frac{E}{T}\right]\right]^d\right\} \tag{9.14a}$$

and a contribution to the specific heat given by

$$C_V(T)\equiv T\gamma(T)=T\frac{d}{dT}n_I(T)\ . \tag{9.14b}$$

It is clear that both $\chi_{\rm LM}$ and γ diverge as $T\to 0$. The conclusion is that disordered Fermi systems with local moments are not Fermi liquids in any dimension at any finite disorder.

Several points should be stressed. First, the above argument is of the Lifshitz-tail type and presumably is correct only asymptotically. The above reasoning therefore might be valid only at extremely low temperatures. Second, the argument does not depend at all on the metal-insulator transition or on any long-wavelength physics. Third, we have not yet considered the effects of local-moment/local-moment interactions. It can be argued that the latter themselves can lead to important quenching if the range of the interactions is long enough.

Let us examine the final point above (Bhatt and Fisher, 1992). In the metallic phase the interaction between local moments occurs via the Ruderman-Kittel-Kasuya-Yosida (RKKY) interaction mediated by the conduction electrons. The interaction between spins at positions $\mathbf{R}_i$ and $\mathbf{R}_j$, separated by a distance R_{ij} and coupled to the conduction electrons via exchanges J_i and J_j, is

$$K_{ij}\sim\frac{J_iJ_j}{|R_{ij}|^d}N_Fg_d(k_F\mathbf{R}_i,k_F\mathbf{R}_j)\ , \tag{9.15}$$

with g_d an oscillating function, $|g_d|=O(1)$. For spin-$\frac{1}{2}$ local moments, the antiferromagnetic interactions have a stronger effect than the ferromagnetic interactions. For simplicity, g_d in Eq. (9.15) is replaced by unity. To estimate the effects of Eq. (9.15) we consider two local moments in cavities of size r and separated by a distance R. The typical separation R can be related to $n_I(r)$ by $R\sim n_I^{-1/d}$. To relate R to the temperature it is assumed that two spins will form a nonmagnetic singlet state if the energy gained is greater than T. This gives

$$T\sim\frac{J^2(r)}{R^d}N_F \tag{9.16a}$$

or

$$n_I\sim\frac{T}{N_FJ^2(r)}\ . \tag{9.16b}$$

Equations (9.16), (9.10), and $n_I\sim e^{-cr^d}$ lead to a low-temperature local-moment susceptibility given by

$$\chi_{\rm LM}(T)\sim\frac{1}{N_FJ^2(r)}\sim\exp\{c[\ln(1/T)]^{1/d}\}\ . \tag{9.17a}$$

This and a similar expression for $\gamma(T)$ diverge slower than any power of T as $T\to 0$. Note that this estimate should be considered as a (nonrigorous) lower bound for the susceptibility. It is based on classical reasoning, which can underestimate singularities caused by quantum effects. Also note that the amplitude of the RKKY interaction must vanish (perhaps related to the vanishing of the single-particle DOS) at the metal-insulator transition. This would lead to a singularity in $\chi_{\rm LM}$ that is controlled by the amplitude of K_{ij}.

Finally, we compare the local-moment quenching discussed above with that of Sec. IX.B.1, where a short-ranged interaction between local moments was assumed. Suppose $K_{ij}\sim|R_{ij}|^{-a}$. Repeating the arguments given above yields

$$\chi_{\rm LM}(T)\sim\frac{1}{T^{1-d/a}}\exp\left[\frac{c}{a}[\ln(1/T)]^{1/d}\right]\ , \tag{9.17b}$$

i.e., $a=d$ is a marginal case. Also note that for exponential interactions a is infinite, and the mechanism discussed does not quench the local moment.

Very recently another way to avoid Kondo screening was proposed by Dobrosavljevic, Kirkpatrick, and Kotliar (1992). This mechanism is strongest near (but not at) the metal-insulator transition and at temperatures that are not too low. These authors argue that, in contrast to the ideas explained above, in which the distribution of the Kondo temperature is due to the distribution of J's in Eq. (9.9), a similar effect can be obtained by using the fact that the single-particle DOS in Eq. (9.9) is a local DOS, $N_F=N_F(r)$, which also has a broad distribution. In particular it is known that away from the metal-insulator transition the distribution $\mathcal{P}[N_F(r)]$ is close to Gaussian (but with log-normal asymptotics), while near the transition it becomes completely log-normal. Let $u=\ln(\sigma_0/\sigma)$ with σ_0 the bare conductivity and σ its renormalized value. For $u\geq 1$ one has (Lerner, 1988)

$$\mathcal{P}[N_F(r)]=\frac{1}{2\sqrt{\pi u}\,N_F(r)}\exp\left[-\frac{1}{4u}\ln^2\left[\frac{N_F(r)}{N_F}e^u\right]\right]\ , \tag{9.18}$$

with N_F the average DOS. In the tails of the distribution, Eq. (9.18) holds for all u. With Eq. (9.9) this leads to a distribution of Kondo temperatures given by

$$\mathcal{P}(T_K)=\frac{1}{2\sqrt{\pi u}\,T_K\ln(E/T_K)}\times\exp\left[-\frac{1}{4u}\ln^2\left\{\frac{N_FJ}{e^u}\left[\ln\frac{E}{T_K}\right]\right\}\right]\ . \tag{9.19}$$

We note that the distribution $\mathcal{P}(T_K)$ diverges as $T_K\to 0$, which implies that there is a large number of spins with a very low Kondo temperature. This distribution can be used to compute the low-temperature form of χ and γ. A local moment contributes a term $\sim 1/T$ to both χ and γ if $T_K<T$ at that site. Thus we can write for χ (and similarly for γ)

$$\chi_{\rm LM} \sim \frac{1}{T} n_{\rm free}(T) \ , \tag{9.20a}$$

where we have defined an effective number of free spins by

$$n_{\rm free}(T) = \int_0^T dT_K \mathcal{P}(T_K) \ . \tag{9.20b}$$

Introducing the quantity

$$x_{\rm max}(T) = -\frac{1}{\sqrt{4u}} \ln[N_F J e^{-u} \ln(E/T)] \ , \tag{9.20c}$$

we can write

$$n_{\rm free}(T) = \frac{1}{\sqrt{\mu}} \int_{-\infty}^{x_{\rm max}(T)} dx \, \exp[-x^2]$$

$$= \tfrac{1}{2}\{1 + {\rm erf}[x_{\rm max}(T)]\} \ . \tag{9.20d}$$

At low temperatures, $n_{\rm free}(T) \sim T^{\alpha(T)}$, with $\alpha(T \to 0) \to 0$, i.e., $n_{\rm free}$ vanishes, but slower than any power of T. With Eq. (9.20a) we conclude that the distribution of $N_F(r)$ leads to a divergent $\chi_{\rm LM}$ (and γ) as $T \to 0$.

It is important to estimate the magnitude of the anomalies predicted by Eqs. (9.20) for typical parameter values. With $E \sim E_F \sim 10^4$ K, a bare Kondo temperature of about 10^2 K, and in the strongly disordered (but still metallic) region where $u \cong 1$, one finds $n_{\rm free}(T) \sim 40$–$60\,\%$ at $T \sim 10^{-4}$–1 K. Note that the extremely slow temperature dependence of $n_{\rm free}$ is a consequence of the $\ln\ln(T^{-1})$ dependence of $x_{\rm max}(T)$.

The preceding argument neglects the interaction between local moments. Repeating the discussion below Eq. (9.15) leads to only a logarithmic singularity in $\chi(T \to 0)$.

We next compare the theoretical results above with the experimental ones given in Figs.25 and 39. Both theory and experiment suggest a low-temperature singularity in the magnetic susceptibility, even well inside the metallic phase. However, the experiments indicate a power-law singularity, $\chi(T \to 0) \sim T^{-a}$, with $a \cong 0.5$, while Eq. (9.17a) shows a singularity that is weaker than any power law. There are a number of possible explanations for this discrepancy. (1) The experiments are not yet in the asymptotic low-temperature regime. (2) In addition to the local-moment contribution to χ, some of the systems might show magnetic anomalies because they actually are in the incompletely frozen spin phase; see Sec. VI. (3) The basic physical point in the second argument for $T_K = 0$ is that some local moments do not see the conduction electrons and therefore cannot be Kondo screened by them. Following this line of reasoning one can argue that the surviving local moments will not interact through a simple RKKY interaction either, since this interaction is indirectly caused by the conduction electrons. The conclusion is that the current understanding of the interaction between local moments is not sufficient to compare theory and experiment quantitatively.

It is also important to examine how the local moments that survive influence the transport behavior near the metal-insulator transition (see, for example, Sachdev, 1989). Naively one might expect that the magnetic impurity universality class (MI) discussed in Sec. V.A.1 provides the relevant behavior. Upon reflection, however, the situation is more complicated. There are two physical points. First, the number of local moments vanishes at $T \to 0$, which is not reflected in the magnetic impurity universality class. Second, as discussed above, the local moments that survive are probably *not* interacting with the conduction electrons. To discuss this final point in more detail we examine whether or not the surviving local moments can give a mass to the spin-triplet particle-hole propagator that plays a crucial role in the generic universality class discussed in Sec. VI, and which is absent in the magnetic impurity universality class. A related question is how the local moments affect the Cooperons that lead to the logarithmic corrections to scaling discussed in Sec. V.A.3.b.

At the level of the Born approximation for local-moment/electron scattering, the scattering rate for a given site is proportional to the local DOS, $N_F(r)$. Adding the contribution from all the sites with $T_K < T$, and taking into account that the scattering rate is proportional to the DOS, leads one to define an effective fraction of local moments that induce spin-flip scattering of the conduction electrons,

$$n_{\rm sc}(T) = \frac{1}{N_F} \int_0^T dT_K N_F(T_K) \mathcal{P}(T_K) \ . \tag{9.21a}$$

This integral can be evaluated as before, and one finds

$$n_{\rm sc}(T) = \tfrac{1}{2}\{1 + {\rm erf}[x_{\rm max}(T) - \sqrt{u}\,]\} \ . \tag{9.21b}$$

It is particularly interesting to examine $n_{\rm sc}(T)$ as the metal-insulator transition is approached. Since $x_{\rm max} \cong \sqrt{u}/2$ for $u \gg 1$, the argument of the error function is roughly $-\sqrt{u}/2 \to -\infty$. This yields $n_{\rm sc}(T) \to 0$ at the metal-insulator transition. The local moments seem to decouple from the conduction electrons as the transition is approached. The implication for the universality class when the mass of the triplet propagator vanishes continuously at the metal-insulator transition remains to be worked out. Similarly, the effects of the local moments on the Cooperons have not been studied.

3. Formation of local moments

Local-moment formation in disordered metals has been studied in an approximation where disorder is treated exactly (using numerical methods), while the interactions are taken into account only in a Hartree-Fock or mean-field-like approximation (Milovanovic, Sachdev, and Bhatt, 1989). The calculation starts with a disordered Anderson-Hubbard model,

$$\hat{H} = -\sum_{\substack{i \neq j \\ \sigma}} t_{ij} \hat{a}_{i\sigma}^{+} \hat{a}_{j\sigma} + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow} + \sum_{i,\sigma} (\varepsilon_i - \mu) \hat{a}_{i\sigma}^{+} \hat{a}_{i\sigma} \ , \tag{9.22}$$

where i,j extend over all the sites in the system, $\hat{n}_{i\sigma}=\hat{a}_{i\sigma}^{+}\hat{a}_{i\sigma}$, and both the hopping matrix elements t_{ij} and the on-site energies ε_i are random variables. Physical intuition suggests that fluctuations of the t_{ij} and the ε_i produce local environments favoring the formation of local moments, and the main problem is to find a theoretical criterion for this to occur.

Theories of local-moment formation in clean systems have relied on the following three observations: (i) A Hartree-Fock calculation is adequate to determine the boundary between the local-moment and Fermi-liquid regimes (Haldane, 1978; Krishnamurthy *et al.*, 1980a, 1980b). (ii) The electron local-moment state is associated with a resonance near the Fermi level in the Hartree-Fock Hamiltonian. This quasiparticle eigenstate is not localized (Anderson, 1961; Wolff, 1961). (iii) The interactions between the generated local moments can be ignored. The arguments of the previous subsection indicate this might not be correct in disordered systems.

The Hartree-Fock or mean-field approach to the disordered local-moment formation problem proceeds by finding the single-particle Hamiltonian $\hat{H}_{\rm eff}$ which best approximates the properties of the interacting system,

$$\hat{H}_{\rm eff}(\tilde{\varepsilon}_i,\mathbf{h}_i)=-\sum_{i,j,\sigma} t_{ij}\hat{a}_{i\sigma}^{+}\hat{a}_{j\sigma} +\sum_{i,\sigma}(\tilde{\varepsilon}_i-\mu)\hat{a}_{i\sigma}^{+}\hat{a}_{i\sigma}+\sum_i \mathbf{h}_i\cdot\hat{\mathbf{s}}_i \ , \tag{9.23}$$

where $\hat{\mathbf{s}}_i=\sum_{\alpha,\beta}\hat{a}_{i\alpha}^{+}\vec{\sigma}_{\alpha\beta}\hat{a}_{i\beta}/2$ is the spin of the electron at site i. The variational parameters in $\hat{H}_{\rm eff}$ are $\tilde{\varepsilon}_i$ and $\mathbf{h}_i$, representing the local site energy and the local magnetic field, respectively. The best variational wave functions of $\hat{H}_{\rm eff}$ are given by minimizing the free energy $F_{\rm eff}$ with respect to $\tilde{\varepsilon}_i$ and $\mathbf{h}_i$. Since $\mathbf{h}_i=0$ at high T, it is convenient to expand in powers of $\mathbf{h}_i$. Let γ_α be the eigenvalues, and $\Psi_\alpha(i)$ the corresponding wave functions, of $\hat{H}_{\rm eff}$ with $\mathbf{h}_i=0$. The condition that $F_{\rm eff}(\tilde{\varepsilon}_i,\mathbf{h}_i=0)$ has a local minimum with respect to $\tilde{\varepsilon}_i$ yields the self-consistency equation,

$$\tilde{\varepsilon}_i=\varepsilon_i+U\sum_\alpha |\Psi_\alpha(i)|^2 f(\gamma_\alpha) \ , \tag{9.24}$$

where f is the Fermi function. This equation gives $\tilde{\varepsilon}_i$. Expanding $F_{\rm eff}$ to second order in $\mathbf{h}_i$ gives

$$F_{\rm eff}(\tilde{\varepsilon}_i,\mathbf{h}_i)=F_{\rm eff}(\tilde{\varepsilon}_i,\mathbf{h}_i=0) +\sum_{i,j,k}\frac{\chi_{ij}}{4}(\delta_{jk}-U\chi_{jk})(\mathbf{h}_i\cdot\mathbf{h}_k)+O(\mathbf{h}_i^4) \ , \tag{9.25a}$$

where

$$\chi_{ij}=-\sum_{\alpha,\beta}\Psi_\alpha(i)\Psi_\beta^{*}(i)\Psi_\alpha^{*}(j)\Psi_\beta(j)\frac{f(\gamma_\alpha)-f(\gamma_\beta)}{\gamma_\alpha-\gamma_\beta} \tag{9.25b}$$

is the spin susceptibility of free electrons described by $\hat{H}_{\rm eff}$. The eigenfunctions of χ_{ij} will be denoted by $m_\alpha(i)$ and the corresponding eigenvalues by κ_α:

$$\sum_i \chi_{ij}m_\alpha(j)=\kappa_\alpha m_\alpha(i) \ . \tag{9.26}$$

It is clear from Eq. (9.25b) that $\mathbf{h}_i\neq 0$ will first be energetically favorable when $\max(\kappa_\alpha)>1/U$. This is the theoretical criterion mentioned below Eq. (9.22).

The properties of the single-particle wave functions $\Psi_\alpha(i)$ and the eigenvectors $m_\alpha(i)$ of χ_{ij} have been studied using parameters appropriate for Si:P. In general one finds that every $m_\alpha(i)$ that is localized around a site $\mathbf{r}_k$ is associated with an eigenvector $\Psi_\alpha(i)$, which is peaked at $\mathbf{r}_k$ and has an energy close to the Fermi level. To get a quantitative measure of the localization properties of $\Psi_\alpha(i)$ and $m_\alpha(i)$, Milovanovic *et al.* have computed the inverse participation ratios,

$$P_{H\alpha}=\sum_i |\Psi_\alpha(i)|^4 \ , \tag{9.27a}$$

$$P_{\chi\alpha}=\sum_i |m_\alpha(i)|^4 \ . \tag{9.27b}$$

In Fig. 55 the average $P_H=\langle P_{H\alpha}\rangle$ over many samples and for states within an energy $0.1t_0$ (t_0 setting the scale for the t_{ij}) of the Fermi energy is shown. On a log-log plot P_H decreases monotonically as a function of the number of sites in the system, which is evidence that all the states within $0.1t_0$ of the Fermi level are extended. For the same samples, the behavior of $P_\chi=\langle P_{\chi\alpha}\rangle$ is quite different. It remains independent of the system size, showing that while the electronic states are delocalized there exist local-moment states that are localized, i.e., local moments exist at low temperatures in the metallic state.

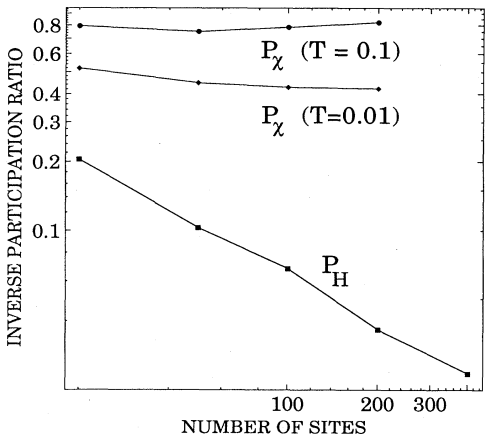

FIG. 55. Mean inverse participation ratio's P_H and P_χ for the eigenvectors $\Psi_\alpha(i)$ of the Hartree-Fock Hamiltonian and for the eigenvectors $m_\alpha(i)$ of χ, respectively, at a density $na^3=0.02$ and filling factor $\frac{1}{2}$ for different system sizes. The values of P_χ are shown at two temperatures. These results are independent of system size, suggesting that the eigenvectors $m_\alpha(i)$ are localized. Values of P_H, on the other hand, decrease with increasing system size, suggesting that the $\Psi_\alpha(i)$ are extended. From Milovanovic *et al.* (1989).

Figure 55 is for an electron filling factor of one electron per site, which is appropriate for uncompensated semiconductors. To investigate the consequences of changing the filling factor, the chemical potential has been varied so that the filling factor decreases to 0.25. In Fig. 56 the results for $F_{\rm LM}$, the ratio of the number of local-moment instabilities to the total number of electrons as a function of filling factor, are shown. One sees that $F_{\rm LM}$ has a maximum at half filling and gradually falls off with decreasing filling factor. This prediction is in disagreement with recent experiments. Hirsch *et al.* (1992) have found that, if one wants to interpret the magnetic properties of doped Si in terms of local moments, $F_{\rm LM}$ in the compensated material Si:P,B is three to five times larger than in uncompensated Si:P.

It is clear from the solution method (expansion in $\mathbf{h}_i$) that the variational treatment used here neglects the interaction between local moments. The arguments reviewed in the previous subsection indicate that this might not be valid. It is also possible that the considerations given here are correct for a range of temperatures, but that at a lower temperature the importance of local moments is reduced for reasons outlined previously. One should also keep in mind that a Hartree-Fock approach often shows an instability that actually is suppressed by more complicated many-body effects. Very recently Dobrosavljevic and Kotliar (1993) have studied local-moment formation in a disordered Hubbard model that is exactly soluble in the limit of infinite dimension. The basic idea is that local-moment formation is a local phenomenon, which should be qualitatively the same in all dimensions, so that results for $d\to\infty$ might be relevant for $d=3$. We note that this would not be the case for the metal-insulator transition, which is presumably controlled by long-wavelength physics. The results of these authors are in general consistent with those of Milovanovic *et al.*, and, insofar as the $d=\infty$ results are valid in $d=3$, justify the Hartree-Fock approximation used above.

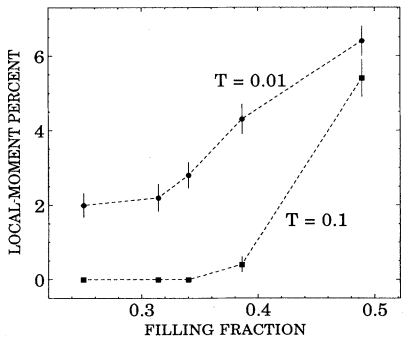

FIG. 56. Ratio of the number of local moments in Hartree-Fock approximation to the total number of electrons as a function of filling factor for $na^3=0.02$ at two different temperatures. From Milovanovic *et al.* (1989).

4. Two-fluid model

The experimental data on transport and thermodynamic properties near the metal-insulator transition have been described by postulating a two-fluid model which imagines a disordered electron system as consisting of a mixture of itinerant electrons and local moments (Quirt and Marko, 1971; Ue and Maekawa, 1971; Paalanen, Graebner, Bhatt, and Sachdev, 1988). This picture has been buttressed by the theoretical developments discussed in the previous two subsections.

The two-fluid model assumes that the results of Bhatt and Lee for the insulating side of the metal-insulator transition, which were discussed in Sec. IX.B.1 above, also hold for the local-moment component of the two-component system on the metallic side. The numerical RG results of Bhatt and Lee can be fit by

$$\frac{\chi(T)}{\chi_0}=\beta\left[\frac{T}{T_0}\right]^{-\alpha}, \tag{9.28a}$$

$$\frac{\gamma(T)}{\gamma_0}=\left[\frac{T}{T_0}\right]^{-\alpha}, \tag{9.28b}$$

with $\alpha\cong 0.6$, T_0 a temperature scale, and $\beta\cong 3.1e^{0.4\alpha}/(1-\alpha)^2$.

On the metallic side of the metal-insulator transition, the susceptibility and specific heat are described by adding a Fermi-liquid contribution,

$$\frac{\gamma(T)}{\gamma_0}=\frac{m^*}{m_0^*}+\left[\frac{T}{T_0}\right]^{-\alpha}, \tag{9.29a}$$

$$\frac{\chi(T)}{\chi_0}=\frac{m^*}{m_0^*}+\beta\left[\frac{T}{T_0}\right]^{-\alpha}. \tag{9.29b}$$

For applications to Si:P, m_0^* is taken to be the Si conduction-band mass ($m_0^*=0.34m_0$) and m^* is the Fermi-liquid effective mass, Eq. (3.126d). The exponent α in Eqs. (9.29) is assumed to be independent of the phosphorus density n. The parameter T_0 measures the fraction of electrons with localized spins. An interesting experimental quantity is the Wilson ratio, $w=(\chi/\gamma)/(\chi_0/\gamma_0)$. With Eqs. (9.29) the only free parameter in w is $T_0'=T_0(m^*/m)^{1/\alpha}$. Figure 57 shows w as a function of temperature. The solid lines are best fits using Eqs. (9.29) with $T_0=0.11$ K and 0.03 K for $n/n_c=1.09$ and $n/n_c=1.25$, respectively, with n_c the critical phosphorus density for the metal-insulator transition. With these T_0 values Eqs. (9.29) also give good fits to the susceptibility and specific-heat data separately. Note that the results on the insulator side, $n/n_c=0.78$, are consistent with the Bhatt-Lee numerical result for $T\to 0$. In Fig. 58 we show χ/χ_0 versus γ/γ_0. We see again that the two-fluid model, shown as a solid line, is in good agreement with the experiments. Note that this curve is independent of the adjustable parameter T_0.

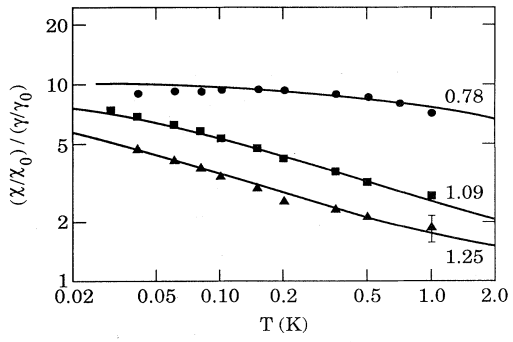

FIG. 57. Ratio of the spin susceptibility to the specific heat in Si:P vs temperature. The solid lines are one-parameter fits with Eqs. (9.29). After Paalanen *et al.* (1988).

Gan and Lee (1986) have invoked a two-fluid model to explain the data of Paalanen, Ruckenstein, and Thomas (1985) on the spin-lattice relaxation time T_1 of ^{29}Si nuclei in Si:P. The experimental results are shown in Figs. 26 and 27. Gan and Lee assume that the relaxation rate consists of a Korringa-type contribution, $1/T_1^K$, and of a local-moment-induced contribution, $1/T_1^{\rm LM}$. $1/T_1^K$ is linear in temperature and independent of the magnetic field B (Korringa, 1950). $1/T_1^{\rm LM}$ is independent of temperature (for reasons we shall see shortly), and its field dependence is determined as follows. Consider a single pair of spins governed by the local-moment Hamiltonian, Eqs. (9.6). The two lowest-lying states are a singlet and a triplet, separated by the exchange coupling J. In a magnetic field, the triplet is split, with the lowest triplet state lowered by $2\mu_e B$, where μ_e is the electronic magnetic moment. When the gap $\Delta_{\rm st}=J-2\mu_e B$ between the singlet and the lowest triplet coincides with the nuclear spin splitting $2\mu_n B$, with μ_n the nuclear magnetic moment, then the spin pair will relax the nuclear spin. Since $\mu_n \ll \mu_e$, the relaxation condition is essentially $J=2\mu_e B$. The relaxation rate is therefore proportional to the probability P of finding the exchange coupling at the right value, $1/T_1 \sim P(J=2\mu_e B)$. Now $P(J)=P_r(r)/|dJ/dr|$, with $P_r(r)$ the probability density for finding a pair separation r. Because of the broad distribution mentioned in Sec. IX.B.1 above, $P_r(r)$ is a slowly varying function of r, while Eq. (9.6b) yields $dJ/dr \sim J$. One therefore has $1/T_1 \sim 1/J \sim 1/B$. This explains the field dependence of the experimentally observed rate, Fig. 27. For the temperature dependence of $1/T_1$ the local-moment mechanism gives a constant contribution, since $T \gg \mu_n B$ for all achievable temperatures. Gan and Lee argue that a constant plus the linear Korringa contribution fits the experimental result, Fig. 26, as well as the power-law fit in the experimental paper, provided only data below $T=100$ mK are considered.

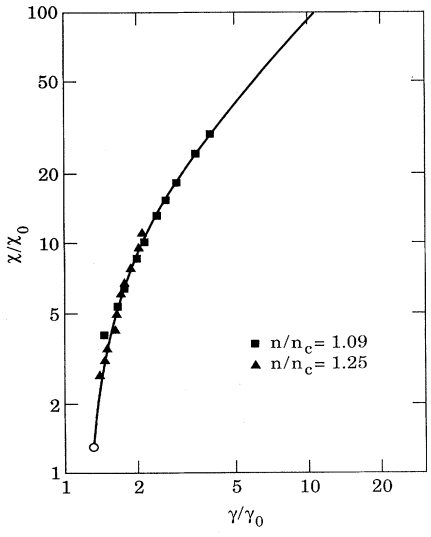

FIG. 58. Susceptibility enhancement vs specific-heat enhancement for metallic Si:P samples. The solid line is the two-fluid model, Eqs. (9.29). After Paalanen *et al.* (1988).

C. Delocalization transition in the quantum Hall effect

Low-temperature magnetotransport experiments on 2-d systems show large regions in the magnetic field B where the Hall resistance ρ_{xy} is accurately quantized according to h/ie^2, with i an integer. This is the integer quantum Hall effect (von Klitzing *et al.*, 1980). In Fig. 59 we show typical experimental results (Wei *et al.*,

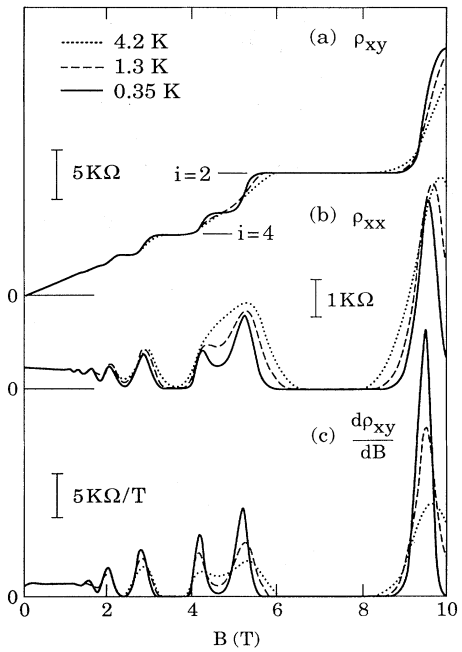

FIG. 59. Quantum transport coefficients in $In_{0.53}Ga_{0.47}As/InP$. (a) ρ_{xy} and (b) ρ_{xx} as functions of magnetic field B at three temperatures, $T=4.2$, 1.3, and 0.35 K. (c) The corresponding $d\rho_{xy}/dB$. The sample had a 2-d electron density $n_{2D}=3.3\times10^{11}$ cm^{-2} and a mobility $\mu=34\,000$ cm^2/Vs at $T=0.8$ K. From Wei *et al.* (1988).

1988) on InGaAs/InP. Note the rather sharp steps connecting the successive quantum Hall plateaus. This is an indication that there is a very narrow range in energy within each Landau level where the electronic states are of a different nature from those in the rest of the band.

In the region where ρ_{xy} changes, the dissipative or parallel resistance ρ_{xx} generally is nonzero. In this region the system behaves like a metal, and extended levels near the Fermi energy are necessary in order to account for the experimental results. The conclusion is that even in 2-d systems there are extended states if the magnetic field is strong enough, and that there are delocalization transitions at critical magnetic fields B^*. Theoretically the number of extended states is expected to be of measure zero, and in some sense the observed transition is probably an insulator-to-insulator transition with a metallic phase of measure zero in between. A sketch of the experimental and theoretical situation is shown in Fig. 60 (Pruisken, 1988).

Motivated by general theoretical considerations like those presented in Sec. IV, as well as by experiments, one finds it natural to assume that this delocalization transition is a conventional zero-temperature phase transition and to construct a scaling theory for it (Pruisken, 1988). The dimensionless distance from the critical point t is given by $\delta B=(B-B^*)/B^*$, and the scaling functions contain an argument $b^{1/\nu}\delta B$, with ν the correlation length exponent. Similarly, the temperature will appear in the combination $b^z T$, with z the dynamical scaling exponent for the delocalization transition. Since the resistance in fundamental units is dimensionless in space dimension 2-d, the scaling part of $\rho_{\alpha\beta}$ ($\alpha,\beta=x,y$) satisfies

$$\rho_{\alpha\beta}(B,T)=\rho_{\alpha\beta}(b^{1/\nu}\delta B, b^z T) . \tag{9.30}$$

Choosing $b=1/T^{1/z}$ gives

$$\rho_{\alpha\beta}(B,T)=\rho_{\alpha\beta}(\delta B/T^{1/\nu z}) . \tag{9.31}$$

In Fig. 61 we show experiments on InGaAs/InP which indicate that the slope of ρ_{xy}, and the width ΔB of the peak in ρ_{xx}, near B^*, satisfy the power laws

$$\left[\frac{\partial\rho_{xy}}{\partial B}\right]_{\max} \sim T^{-\kappa} , \tag{9.32a}$$

$$\Delta B \sim T^{\kappa} , \tag{9.32b}$$

with

$$\kappa=0.42\pm0.04 . \tag{9.33}$$

Equation (9.31) gives

$$\kappa=1/\nu z , \tag{9.34a}$$

and the experimental result yields

$$\nu z \simeq 2.38 . \tag{9.34b}$$

The same scaling behavior with the same value for the exponent κ has been observed for the fractional quantum Hall effect by Engel *et al.* (1990).

These experiments could not separately measure ν or z. Very recently Koch *et al.* (1992) have used a different experimental technique to separately obtain ν and, in prin-

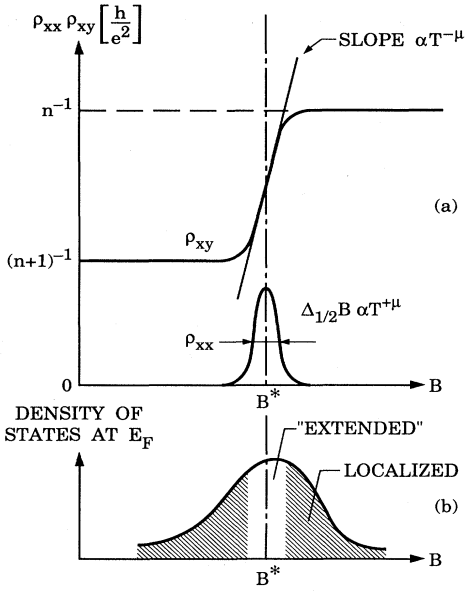

FIG. 60. The resistivity and the density of states in the quantum Hall effect: (a) sketch of the experimentally measured ρ_{xx} and ρ_{xy} as functions of the applied magnetic field B; (b) density of states at the Fermi energy as a function of B. From Pruisken (1988).

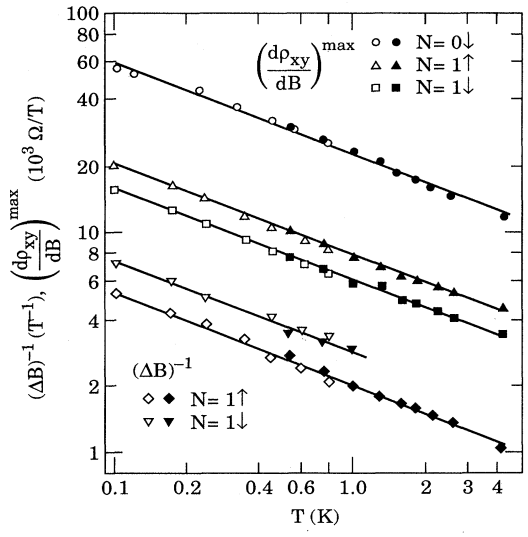

FIG. 61. The upper portion shows the T dependence of $(d\rho_{xy}/dB)_{\max}$ for Landau levels $N=0\downarrow$, $1\uparrow$, and $1\downarrow$; the lower portion shows the T dependence of $1/\Delta B$ for the $N=1\uparrow$ and $1\downarrow$ Landau levels. The open symbols are data taken in a dilution refrigerator, whereas the filled symbols are data taken in a ^{3}He system. The typical uncertainty in T is about 0.02 K at 0.4 K. From Wei *et al.* (1988).

ciple, z. They obtained a universal result for ν, namely, $\nu=2.3\pm0.1$, but for reasons that are not clear, the dynamical scaling exponent did not appear to be universal. There is no theory for either ν or z for interacting systems.

There also has been numerical work on the delocalization transition of noninteracting disordered electrons in a strong magnetic field (Aoki and Ando, 1985; Ando and Aoki, 1985; Huckestein and Kramer, 1990; Liu and Das Sarma, 1993, 1994). The results of these simulations are generally consistent with the above phase-transition scenario. There is a delocalization transition near the center of the Landau level, and the correlation length exponent has been estimated to be on the order of two. The numerical result by Huckestein and Kramer is $\nu=2.34\pm0.04$. Within the error bars this agrees with the experimental result by Koch *et al.* We note that *a priori* there is no reason for the simulation result for a noninteracting system to agree with actual experiment. It is not known whether this agreement is fortuitous or not.

D. Superconductor-insulator transition in two-dimensional films

The theoretical ideas discussed in Sec. IX.A and numerous experimental studies suggest that in 2-*d* films there is a superconductor-insulator (SI) transition at zero temperature, i.e., there exists a quantum phase transition. Theoretically it is easiest to imagine probing such a transition by varying the disorder G. Experimentally, however, it is difficult to vary the disorder in a given sample. Alternatively, a superconductor-insulator transition can be observed for fixed disorder if an external magnetic field B is increased beyond a critical value B_c. In Fig. 62 we show such a transition in amorphous InO_x films (Hebard and Paalanen, 1990). For $B>B_c=5460\pm20G$, the zero-temperature resistance R tends to diverge, and for $B<B_c$ R tends to vanish as $T\to0$. This quantum phase transition appears to be continuous, and it is natural to construct a scaling description of it (Fisher, 1990). Here we restrict ourselves to very simple considerations and refer the reader to the original paper for a more complete theory.

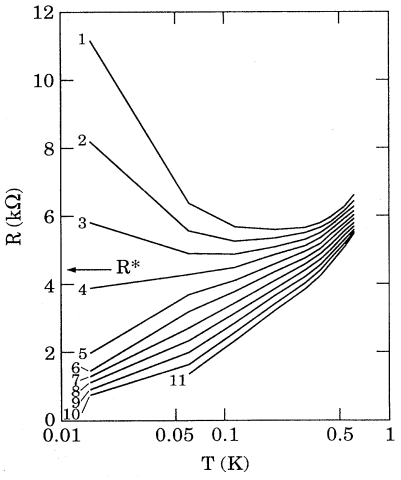

FIG. 62. Logarithmic plots of resistive transitions in a-InO_x. The isomagnetic lines range from $B=4$ KG (11) to $B=6$ KG (1) in steps of 0.2 KG. R^* is the zero-temperature resistance at the critical field. After Hebard and Paalanen (1990).

Let the critical disorder for the superconductor-insulator transition in zero field be G_c. At $G=G_c$, the superconducting transition will be at $T=0$, while for $G<G_c$ there will be a finite-temperature (Kosterlitz-Thouless) superconducting transition with a critical temperature T_c. As $T_c\to0$ its scaling behavior should be determined by the dynamical scaling exponent z for the superconductor-insulator transition. If the correlation length diverges at the transition according to $\xi\sim|G-G_c|^{-\nu}$, then scaling gives

$$T_c\sim\xi^{-z}\sim|G-G_c|^{\nu z}\ . \tag{9.35}$$

Similarly, the critical magnetic field B_c for the field-tuned superconductor-insulator transition should vanish as $G\to G_c$. Since in superconductivity theory, and in the Cooper channel in general, the magnetic field appears as a length (viz., the cyclotron radius) squared, B_c will scale as

$$B_c\sim\xi^{-2}\sim|G-G_c|^{2\nu}\ . \tag{9.36}$$

Equations (9.35) and (9.36) imply

$$B_c\sim T_c^{2/z}\ . \tag{9.37}$$

Finally, by identifying the frequency Ω, or the energy scale, with the long-ranged Coulomb interaction between electrons at scale ξ, which goes as $1/\xi$, it can be argued (Fisher *et al.*, 1990) that

$$z=1\ . \tag{9.38}$$

This result can also be derived by examining the dynamics of density fluctuations in a system with long-ranged interactions; see Eq. (4.15a). Equations (9.37) and (9.38) give

$$B_c\sim T_c^2\ . \tag{9.39}$$

Equation (9.39) is a nontrivial prediction that can be easily tested experimentally. Figure 63 shows B_c versus T_c in amorphous InO_x (Hebard and Paalanen, 1990). The good agreement with Eq. (9.39) shows that the experiment is at least consistent with the scaling theory for the superconductor-insulator transition.

Finally, general theoretical considerations give one other important result. At the critical disorder, and at zero temperature, a 2-*d* film can be regarded as a metal with a universal conductivity σ (Fisher *et al.*, 1990). In other words, the dimensionless critical resistance for the superconductor-insulator transition should be a universal number, R^*. This result can be proven using ideas similar to those used to prove that amplitude ratios are universal at thermal phase transitions (see Kim and

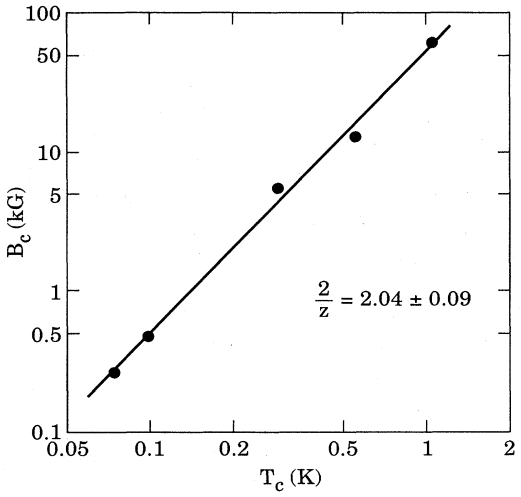

FIG. 63. Logarithmic plot of the critical field B_c vs T_c in a-InO_x for five different films or values of the disorder. From Hebard and Paalanen (1990).

Weichman, 1991). Recent Monte Carlo studies (Sørensen *et al.*, 1992) confirm this idea. So far, experiments have not yet convincingly demonstrated the predicted universality.

X. CONCLUSIONS

In the past few decades there has been an enormous amount of work on the metal-insulator transition in particular and on various properties of disordered electronic systems in general. In this review we have concentrated on the interplay between electron-electron interactions and disorder. We have argued that many aspects of this problem are now fairly well understood, and naturally we have concentrated on these aspects. However, much remains to be understood. In this section we conclude our overview by summarizing some of the main theoretical and experimental achievements and by briefly discussing a number of open problems in the field of disordered electronic systems.

A. Summary of results

One of the most important theoretical advances towards our understanding of the metal-insulator transition was the realization that it can be viewed as a continuous quantum phase transition that is characterized by three independent critical exponents (Wegner, 1976; Abrahams *et al.*, 1979). The remaining exponents that determine the singular behavior of physical quantities near the metal-insulator transition are related to the three independent ones by scaling laws. The general scaling picture for the metal-insulator transition has been given in Sec. IV. To complete this scaling description, methods were developed to calculate the critical exponents explicitly. The major breakthrough that made this possible is due to Wegner (1979). For the special case of noninteracting disordered electrons, he showed that the metal-insulator transition problem could be mapped onto an effective field-theoretic problem. The resulting nonlinear-sigma-model field theory was susceptible to a solution by RG methods, and an $\epsilon = d - 2$ expansion for the critical exponents could be developed. Finkel'stein (1983) generalized the model to include the effects of electron-electron interactions. Finkel'stein's action is derived in Sec. III, and the ϵ expansion is performed in Secs. V and VI. It should be noted that the convergence properties of this ϵ expansion are not at all understood; see Sec. X.B below. There we shall also discuss other conceptual problems of the sigma-model approach.

Another important step towards describing the metal-insulator transition for general physical systems was the identification of distinct universality classes. Using methods from either field theory or many-body perturbation theory it was established, first for noninteracting electrons, that there are four main universality classes for the metal-insulator transition (Efetov *et al.*, 1980): (i) systems with magnetic impurities (class MI), (ii) systems in an external magnetic field (class MF), (iii) systems with strong spin-orbit scattering (class SO), and (iv) systems that have no symmetry-reducing fields, which we call the generic universality class (class G). After including electron-electron interactions, it was realized that one must also distinguish between the cases of short-ranged and long-ranged interactions (Castellani, DiCastro, Lee, and Ma, 1984). The conclusion is that there are at least eight universality classes for the localization transition in disordered interacting fermion systems. These eight universality classes have been listed in Table I, and the respective three independent critical exponents are given in Table III. We remark in passing that one of the main experimental problems is the determination of which universality class is the relevant one for a given system.

Finkel'stein (1984a), Castellani, DiCastro, Lee, and Ma (1984), and Castellani, DiCastro, Forgacs, and Sorella (1984) found that for the magnetic impurity, magnetic field, and spin-orbit universality classes, straightforward perturbative RG methods lead to *bona fide* metal-insulator transitions. The behavior within these universality classes is therefore in principle understood, although, as mentioned above, the unknown (and presumably very poor) convergence properties of the ϵ expansion about the lower critical dimension make a determination of the critical exponents in the physical dimension $d = 3$ impossible. For the generic universality class the situation turned out to be more complex. In the field-theoretic description, a magnetic instability competes with the metal-insulator transition for $\epsilon \ll 1$ (Finkel'stein, 1984c; Castellani and DiCastro, 1986; Belitz and Kirkpatrick, 1989b). It was subsequently shown (Kirkpatrick and Belitz, 1990b) that for small enough ϵ the field theory predicts a phase transition from a Fermi liquid to a magnetic phase. This transition precedes the metal-insulator transition. It was later argued (Kirkpa-

trick and Belitz, 1992a) that in $d=3$ there are some regions of parameter space where these instabilities do not compete, and that there is a direct transition from a Fermi liquid to an insulator. However, a controlled description of this metal-insulator transition has not yet been given. Very recently another prediction has been made (Kirkpatrick and Belitz, 1993, 1994): for the spin-orbit and generic universality classes it has been proposed that there are logarithmic corrections to scaling and to some asymptotic power laws near the metal-insulator transition.

On the experimental side, an important development began with the finding that there are logarithmic corrections to the electrical conductivity in 2-d systems (Dolan and Osheroff, 1979). This result confirmed the theoretical notion that the lower critical dimension for the metal-insulator transition is two. The sign and amplitude of the logarithmic correction was found to depend on whether or not there are magnetic impurities, or spin-orbit scattering, or an external magnetic field present (see, for example, Bergmann, 1984). This is consistent with the theoretical notion of distinct universality classes for the metal-insulator transition that are characterized by the presence of these various symmetry breakers.

Later experiments investigated how the bulk conductivity vanishes at the metal-insulator transition, either as a function of temperature at the critical disorder or as a function of disorder at (extrapolated) zero temperature. The most accurate experiments (Rosenbaum *et al.*, 1983; Dai *et al.*, 1992) find the conductivity exponent at zero temperature, s, to be close to either one-half or unity. Less accurate experiments have reported values that range between one-half and one. These experimental results seemed to be at odds with the theoretical results described above in at least two ways. First, as described in Secs. IV.A.2 and V.A.3.b, any value of s smaller than $\frac{2}{3}$, taken at face value, seems to violate a rigorous lower bound on the correlation length exponent. Second, there are experiments on different materials which should be in the same universality class but are found to have different values of s. An example is Si:B ($s=0.65^{+0.05}_{-0.14}$; Dai *et al.*, 1992), and $Si_{1-x}Au_x$ ($s=0.95\pm0.20$; Nishida *et al.*, 1985), which are both expected to be in the spin-orbit universality class. In Secs. V.B and VI.B both of these problems are tentatively explained as being due to the logarithmic corrections to scaling that exist for the spin-orbit and generic universality classes. However, recently the validity of some of the experimental conclusions has been debated (Stupp *et al.*, 1994a, 1994b; Rosenbaum, Thomas, and Paalanen, 1994). Further investigations of this problem are clearly needed.

Important experiments have also been done on the thermodynamic properties near the metal-insulator transition (Paalanen *et al.*, 1986, 1988). In all systems that were investigated an increasing magnetic susceptibility and specific-heat coefficient were observed in the metallic phase down to the lowest measurable temperatures. This has been interpreted as evidence for the existence, in the metallic phase, of local moments that are not Kondo quenched and survive down to zero temperature. This is a surprising result, which implies that at least in some classes of disordered electronic systems there might not be a Fermi-liquid phase anywhere in the phase diagram.

In this review we have also touched upon the broader subject of disordered electronic systems in general. Section VII was devoted to the effects of disorder on the mean-field superconducting transition temperature in conventional bulk superconductors. According to present understanding, there are four effects that occur in all systems. Denoting the mean-field superconducting transition temperature by T_c, the four effects are (i) A decrease in effective electron-phonon coupling due to longitudinal phonons. This is overcome by an increase (from zero) in the electronic coupling to transverse phonons. The net result is a T_c-increasing mechanism (Keck and Schmid, 1976). (ii) A decrease in the single-particle density of states, which in turn decreases T_c (Maekawa and Fukuyama, 1982). (iii) An increase in the repulsive Coulomb pseudopotential, which also decreases T_c (Anderson *et al.*, 1983). (iv) Either an increase or a decrease of T_c from the disorder dependence of the quasiparticle density of states (Kirkpatrick and Belitz, 1992c). For small disorder the above competing effects can lead to a T_c that either increases or decreases with disorder (Belitz, 1987a). However, for sufficiently strong disorder T_c always decreases and vanishes in the metallic phase. In general the theory compares well with experimental results. There are, however, special problems with 2-d superconducting systems and questions about whether the processes listed above all contribute in $d=2$. This requires further investigations. In Sec. VIII it was shown that in some parts of parameter space and for low enough dimensions, disorder can actually cause a novel type of even-parity spin-triplet superconductivity (Kirkpatrick and Belitz, 1991; Belitz and Kirkpatrick, 1992). This new phase of disordered electronic systems has so far not been observed experimentally.

Finally, there have been important theoretical developments outside of the framework of the nonlinear sigma model, in particular, work on local moments in disordered itinerant electronic systems. It has been argued that (i) fluctuations in local energies can produce local environments favoring the formation of local moments in disordered systems (Milovanovic *et al.*, 1989); (ii) rare fluctuations in either the conduction-electron/local-moment coupling constant (Bhatt and Fisher, 1992), or in the local single-particle density of states (Dobrosavljevic *et al.*, 1992) can effectively cause the Kondo temperature to be zero in disordered electronic systems. These theoretical results were motivated by and are consistent with experiments in doped semiconductors such as Si:P, Si:P,B, and Si:B. They have been described in Sec. IX.B.

B. Open problems

We conclude by mentioning a number of open problems related to disordered electronic systems. Our re-

marks and questions are necessarily speculative, and the order in which they are arranged is to some extent arbitrary.

1. Problems concerning local moments

In Secs. V.B, VI.B, and IX.B it was argued that there is both experimental and theoretical evidence for the existence of local moments in the metallic phase of doped semiconductors at zero temperature. While some work has been done addressing the formation of, and the absence of Kondo screening for, these local moments, an important unresolved question is how they interact with the conduction electrons that become localized at the metal-insulator transition. The answer is of vital importance for the interpretation of transport experiments. It would be contained in a solution of the very difficult problem of constructing a theory for the metal-insulator transition that self-consistently takes into account the disordered conduction electrons, local-moment formation and Kondo physics in disordered systems, and local-moment interactions. In particular, it would be important to know how local moments affect the strong spin-density fluctuations described in Sec. VI and the logarithmic corrections to scaling discussed in Sec. V.A.3.b.

The above remarks bring us to an important point. In Sec. IX.B it was argued (Bhatt and Fisher, 1992) that a generic disordered electronic system might not possess a Fermi-liquid phase. Naively, this notion casts doubt on the validity of the sigma model presented in Sec. III, in the derivation of which a disordered Fermi-liquid description was assumed. However, from a more general point of view, the sigma-model approach relies more on fundamental conservation laws than it does on the validity of a Fermi-liquid description. The key physical idea is to determine the slow modes of the system and to construct an effective theory for them. In general the basic slow modes are determined by the conservation laws for charge, energy, and spin, independent of whether or not the ground state is a Fermi liquid. From this viewpoint the concluding questions in the paragraph above are more important than whether or not a particular system has a Fermi-liquid phase. A question of particular relevance for the issues discussed in Secs. V.A.3 and VI is whether the local moments act like magnetic impurities and lead to spin-flip scattering. If this were the case, then spin diffusion would always be eliminated as a slow mode, and neither the generic nor the spin-orbit universality class would be realized in these systems. As was discussed in Sec. IX.B.2, there are indications that this is not the case, and that the interaction between local moments and conduction electrons is weak or effectively absent, but more work on this point is needed. Some relevant experimental information on this point was recently provided by Allen, Paalanen, and Bhatt (1993), who conclude from studies of NbSi films that the local-moment density is not correlated with the effective critical exponent for the conductivity at the metal-insulator transition. However, recent computer simulations cast some doubt on the conclusiveness of this study (Bhatt, 1993).

From a general point of view it is also possible that in some systems a non-Fermi-liquid metallic phase exists with long-range correlations that are not related to the underlying conservation laws. In this case the present sigma-model approach might be invalid because these additional slow modes could affect the metal-insulator transition. In this context it is interesting that recent work on the extended Hubbard model in infinite dimensions has established that, in some parts of parameter space, there is indeed a non-Fermi-liquid metallic phase with correlation functions that are characterized by power-law decays (Si and Kotliar, 1993).

2. Description of known or expected phase transitions

The discussion in Secs. VI and VII was incomplete because for various phases and phase transitions in the phase diagram that are known or suspected to exist there is not yet a satisfactory description.

a. Phase diagram for the generic universality class

In Sec. VI it was argued that for the generic universality class there is a direct Fermi-liquid–insulator transition if $\epsilon = d - 2$ is not too small. It was pointed out that this transition is inaccessible within the framework of the ϵ expansion and therefore it is nonperturbative in nature. In Sec. VI a tentative description of this metal-insulator transition was discussed, which is, however, not a controlled description in the usual sense because there is no small parameter. Even if we restrict ourselves to the sigma-model description of disordered electronic systems, it is an open problem to confirm systematically the phase diagram given in Fig. 28 and to develop a way to describe the transition in a controlled manner.

In particular, it should be pointed out that the existence of the multicritical point in Fig. 28 has not been proven. Rather its existence was suggested by solving equations that were derived for large γ_t. For initial conditions γ_t^0 smaller than some $(\gamma_t^0)_c$ these equations were found not to have a solution with $\gamma_t \to \infty$ self-consistently. This was interpreted as an indication that there is no magnetic incompletely frozen spin phase between the Fermi-liquid and insulator phases. However, this is not the only possible interpretation. It is conceivable that the IFS phase exists for all γ_t^0, but that one should start using the equations derived in Sec. VI.A only when γ_t has already been renormalized to a large value. At present we cannot rule out this possibility. It should be emphasized, however, that the direct Fermi-liquid-to-insulator phase transition described in Sec. VI.A.3 was one of the generic possibilities discussed in Sec. IV.A and on general grounds one would expect it to occur somewhere in parameter space.

Also, the physical nature of the IFS phase needs to be clarified. The existing interpretation as a critical phase for the spin dynamics is a speculative one, based on the properties of integral equations that are not valid in the new phase. Other possibilities, including that of long-range spin order, cannot be ruled out. Again, the nonlinear sigma model may not be well suited for this task, since the description of any large-disorder phase is very difficult within this approach. The expected transition from the IFS phase to the charge insulator can obviously be studied only after this problem has been solved.

b. Critical-exponent problems

In Secs. V and VI we discussed the conductivity exponent puzzle, i.e., the fact that the observed critical behavior of the conductivity appears to be inconsistent with a rigorous bound on the exponent ν. This has been tentatively explained as due to logarithmic corrections to scaling; see Fig. 38. However, other explanations are conceivable. For instance, the lower bound on ν would be inapplicable if the transition were unconventional in the sense that there was more than one divergent length scale. This can happen in disordered systems; an example was studied recently by Fisher (1992). Also, Wegner scaling would be violated, and the conductivity exponent s would be different from ν in $d=3$ if the disorder scaled to zero at the transition, as was first suggested by Castellani, Kotliar, and Lee (1987). A renormalization of the charge would have the same effect (Belitz and Kirkpatrick, 1993). As we have discussed in Secs. V and VI, these possibilities have been investigated in some detail and found not to be realized in any of the known universality classes. However, this statement holds only within the framework of the sigma-model approach. It is possible that the sigma model is not sufficiently general (see Sec. X.B.3 below), and in a more general model this scaling scenario might be realized.

In Sec. V.B.1 we also mentioned another exponent puzzle that deserves closer attention. Measurements of the Hall coefficient seem to show a correlation between the values of the critical exponents s for the conductivity and s_H for the Hall conductivity: for systems with $s \cong 1$ one finds $s_H \cong 1$, while for systems that show the "anomalous" value $s \cong 0.5$ the Hall coefficient is found to be uncritical, i.e, $s_H = 0$. The only reported exceptions to this rule seem to be AlGaAs:Si, where $s=1$ and $s_H = 0$ (Katsumoto *et al.*, 1987, 1989), and Si:B, where $s \cong 0.5$ and $s_H > 0$ (Dai, Zhang, and Sarachik, 1993). On the experimental side it should be pointed out that the existing measurements of the Hall coefficient are not nearly as accurate as the best available data for s. Moreover, noncriticality of an observable is hard to establish experimentally, since the critical region for the observable may be very small. Theoretically, the Hall effect has not been considered for the case of interacting electrons. Clearly this problem needs more work both theoretically and experimentally. Efforts in this direction have been made (Kotliar, 1992; Sarachik, 1992), and hopefully this issue will be resolved soon.

c. The superconductor-metal transition

In Sec. IX the superconductor-insulator transition that presumably takes place at zero temperature for $d \leq 2$ was briefly discussed. In Sec. VII a mean-field theory for T_c degradation in conventional superconductors was reviewed. It was noted that experiments seem to indicate that there is a zero-temperature superconductor-metal transition in bulk superconductors. The nature of this quantum phase transition has not been investigated and is an important open problem. It is not at all clear that the formalism developed in Sec. VII to describe the finite-temperature superconductivity phase transition can be used at zero temperature. The main problem is that the coefficients in the effective Hamiltonian for the order-parameter fluctuations diverge as $T \to 0$. Physically this reflects long-range correlations at zero temperature that are caused by a coupling of statics to dynamics. In general dynamical spin and charge-density fluctuations are of long range because of the conservation laws. One possible solution to the above problem is to *not* integrate out the fluctuations that lead to these nonlocal terms. If noncritical amplitude fluctuations are integrated out, then the resulting field theory is just the nonlinear sigma model discussed in the bulk of this review, except that $K_c^{(s)} < 0$. Alternatively, it might be useful to consider the field theory derived in Sec. III before integrating out the massive modes. A controlled RG solution of either of these field theories has not yet been given.

d. The quantum Hall effect

In Sec. IX.C we discussed experimental and theoretical evidence for scaling and a delocalization transition in the integer quantum Hall effect. While there is a good microscopic understanding of this effect in general (Prange and Girvin, 1987), this is not true for the delocalization transition aspect of it. For instance, it has so far not been possible to extract this transition from the field theory for the integer quantum Hall effect (Pruisken, 1984) or to calculate the relevant exponents. Further, no attempts have been made to incorporate the electron-electron interaction into the theory of the integer quantum Hall effect and to arrive in this way at a unified theory of both the integer and the fractional quantum Hall effects. There are, however, developments towards such a unification that start from the fractional quantum Hall effect (see, for example, Jain, 1992).

3. Problems with the nonlinear sigma model

Throughout much of this review we have used the field-theoretic nonlinear-sigma-model approach to describe disordered electronic systems. Prior to its utiliza-

tion in this field, the nonlinear sigma model had already found condensed-matter applications, for example as a model for Heisenberg ferromagnets (Polyakov, 1975; see also Zinn-Justin, 1989). Despite a considerable amount of work on this model, its status as a model for these condensed-matter problems is not entirely clear. There are several fundamental points that are not understood. Some of these are listed in this section.

a. Nonperturbative aspects and completeness

It is clear that the small-disorder (G) or ε-expansion approach misses terms that are exponentially small as $G \to 0$. For example, the parametrization given by Eq. (3.117a) implies the constraint $1-qq^{+}>0$, which is not enforced in the small-disorder expansion. Technically, one makes the approximation of extending the functional integral over all q. Physically, the fluctuations that violate the above constraint are of finite amplitude and contribute terms of $O[\exp(-1/Gk_F^{d-2})]$ to the observables. It is certainly not clear that these terms can be neglected in $d=3$, where the dimensionless disorder is of order unity at the metal-insulator transition.

Even if the nonlinear sigma model could be solved exactly, it is not obvious that, for example, the O(3) sigma model would give the exact critical behavior of the O(3) Heisenberg model in $d=3$. The equivalence has been shown only in a perturbative sense, and terms which are exponentially small and irrelevant near $d=2$ can become relevant and change the fixed-point structure as d increases (Cardy and Hamber, 1980). In the context of localization it has been noted that the noncompact and compact models that arise from a bosonic (Wegner, 1979) and fermionic (Efetov *et al.*, 1980) formulation of the noninteracting localization problem are perturbatively equivalent, but may differ in their nonperturbative properties. Efetov (1982, 1983) has proposed an alternative, graded or "supersymmetric" nonlinear sigma model for noninteracting electrons that comprises both bosonic and fermionic degrees of freedom and avoids the replica trick. Perturbatively it is equivalent to the replicated models, and its nonperturbative properties have been investigated in the Migdal-Kadanoff approximation (Zirnbauer, 1988). It is currently unclear whether any of these ideas can be used for the interacting case, or what exactly the role of nonperturbative effects is for disordered electrons in general.

A possibly related problem is the poor convergence properties of the $\varepsilon=d-2$ expansion. For the relatively simple cases of noninteracting electrons and the $O(n)$ model, where high-order ε-expansion calculations are possible, this problem has been studied directly in some detail (Hikami and Brézin, 1978; Wegner, 1989; Hikami, 1992; see also Zinn-Justin, 1989, and Sec. III.C.1 above). The conclusion is that the ε expansion is probably not Borel summable for at least some of these models, and reliable estimates for exponents in $d=3$ are hard to obtain from the perturbative renormalization group. For the more complicated interacting electron problem, even high-order ε expansions do not seem feasible using the theoretical formalism described in this review. Further, although for noninteracting electrons numerical techniques have been used to determine the critical exponents (Evangelou, 1990, 1991; Kramer *et al.* 1990), it seems unlikely that similar techniques will be available for the interacting problem in the near future. In summary, there are no practical, controlled ways to calculate critical exponents for the metal-insulator transition in $d=3$.

Another, possibly related, question has been asked with respect to the completeness of the nonlinear sigma model. In deriving the model, gradient terms of higher than second order have been neglected because of their presumed irrelevance. Recent work (Altshuler *et al.*, 1988; Kravtsov *et al.*, 1988; Lerner and Wegner, 1990; Wegner, 1990) has cast doubts on this approximation. The main point is as follows. The dimension of an operator with $2n$ gradients, which to zero-loop order is $2-2n$, making operators with $n>1$ irrelevant, acquires at one-loop order a positive correction proportional to $n(n-1)\varepsilon$. For $n \to \infty$ these corrections are not small for any ε, and they have the potential for overwhelming the negative zeroth-order term. The implications of this observation are not entirely clear. Possible scenarios have been discussed by Wegner (1990), and more work on this problem is desirable.

b. Symmetry considerations and renormalizability

For the nonlinear sigma model describing noninteracting electrons the symmetry properties have been discussed in detail, and renormalizability has been proven (see Zinn-Justin, 1989, and Secs. III.B.4.a and III.C.1 above). This is not the case for the interacting model, Eq. (3.118), which we have discussed in much of this review. While it is clear that the interaction term S_{int} in the action breaks the symmetry of the noninteracting nonlinear sigma model, the nature of the remaining symmetry properties of the full action has not been investigated. Neither has it been proven that the full model is renormalizable. If one neglects the particle-particle channel, all existing (one- and two-loop) calculations indicate that the theory is probably renormalizable with five renormalization constants. However, in general one cannot neglect the particle-particle channel, and in Secs. V [e.g., Eq. (5.38c)], VII, and VIII we have seen that the particle-particle channel contains unusual structures, as far as field theories go, and renormalizability seems far from obvious. It should be emphasized that nonrenormalizability would not necessarily render the theory useless (see, for example, Weinberg, 1983; Collins, 1984, Chap. 5.7.6; Lepage, 1990), but it would be good to know.

4. Alternatives to the nonlinear sigma model

Given the problems with the nonlinear sigma model discussed in the previous subsection, it would be nice to have an alternative approach. In particular, it would be desirable to develop a mean-field theory for the disordered interacting electron metal-insulator transition. For the special case of noninteracting disordered electrons some progress has been made by studying the localization transition on a Bethe lattice (Efetov, 1984, 1987; Zirnbauer, 1986) or in an ensemble of sparse matrices (Fyodorov and Mirlin, 1991). Similar results are obtained by treating the graded or supersymmetric nonlinear sigma model in an effective-medium approximation (Efetov, 1988; Efetov and Viehweger, 1992). The resulting phase transition is rather different from the one expected on the basis of either the $\epsilon = d - 2$ expansion approach or three-dimensional computer simulations. In fact, the localization transition on the Bethe lattice resembles the mean-field spin-glass phase transition. However, it is possible that a more conventional mean-field theory can be constructed for interacting disordered electrons. This is conceivable because, as was discussed in Sec. IV, at the interacting metal-insulator transition the critical exponent β is not fixed to be zero, and therefore it is possible that a finite upper critical dimension exists. Research in this direction might start with the recent work on the Hubbard model in high dimensions (Metzner and Vollhardt, 1989). Alternatively an order-parameter description of the interacting metal-insulator transition might be possible. The unsuccessful attempt of Harris and Lubensky (1981) to give an order-parameter description of the noninteracting metal-insulator transition may give some inspiration in this respect.

5. The 2-*d* ground-state problem

In Sec. VIII the 2-*d* ground-state problem was introduced. The basic point is that if the local moments discussed in Sec. X.B.1 above do not suppress triplet fluctuations in a generic system, then in two dimensions the nature of the ground state is not obvious. Naively the results of Sec. VI would suggest some type of random magnetic insulator, with unknown properties. However, the situation could be more complicated. The results of Sec. VIII indicate that this state will compete with an even-parity spin-triplet superconducting state. Whatever the outcome of that competition is, it seems likely that with decreasing temperature the system will go through a series of regions dominated by different fluctuations. Existing theory is not well suited to deal with the 2-*d* problem, which presents formidable technical difficulties already in the noninteracting case (Houghton, Schäfer, and Wegner, 1980).

ACKNOWLEDGMENTS

Over the years we have benefited from discussions with more colleagues than we can list here individually. While writing this review, conversations with Claudio Castellani, Robert Dynes, John Graybeal, Mikko Paalanen, Thomas Rosenbaum, Myriam Sarachik, and James Valles were especially helpful. This work was supported by the National Science Foundation under Grants Nos. DMR-88-19302, DMR-89-13095, and DMR-92-09879. A portion of this work was done at the Institute for Theoretical Physics at the University of California at Santa Barbara and supported by the National Science Foundation under Grant No. PHY-89-04035. We thank the Institute for hospitality, and all participants in the workshop on Quantum Phase Transitions for stimulating discussions.

REFERENCES

Abeles, B., 1976, in *Applied Solid State Science,* Vol. 6, edited by R. Wolfe (Academic, New York), p. 1.
Abrahams, E., P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan, 1979, Phys. Rev. Lett. **42**, 673.
Abrahams, E., P. W. Anderson, and T. V. Ramakrishnan, 1980, Philos. Mag. B **42**, 827.
Abrahams, E., A. Balatsky, J. R. Schrieffer, and P. B. Allen, 1993, Phys. Rev. B **47**, 513.
Abrahams, E., and P. A. Lee, 1986, Phys. Rev. B **33**, 683.
Abrikosov, A. A., and L. P. Gorkov, 1958, Zh. Eksp. Teor. Fiz. **55**, 1558 [Sov. Phys. JETP **8**, 1090 (1958)].
Abrikosov, A. A., and L. P. Gorkov, 1959, Zh. Eksp. Teor. Fiz. **36**, 319 [Sov. Phys. JETP **9**, 220 (1959)].
Abrikosov, A. A., and L. P. Gorkov, 1960, Zh. Eksp. Teor. Fiz. **39**, 1781 [Sov. Phys. JETP **10**, 593 (1960)].
Abrikosov, A. A., L. P. Gorkov, and I. E. Dzyaloshinski, 1975, *Methods of Quantum Field Theory in Statistical Physics* (Dover, New York).
Abrikosov, A. A., and I. A. Ryzhkin, 1978, Adv. Phys. **27**, 147.
Adams, P. W., D. A. Browne, and M. A. Paalanen, 1992, Phys. Rev. B **45**, 8837.
Ahlheim, U., P. van Aken, H. Spiller, and F. Steglich, 1988, Helv. Phys. Acta **61**, 518.
Alder, B. J., and W. E. Alley, 1978, J. Stat. Phys. **19**, 341.
Alder, B. J., and T. E. Wainwright, 1970, Phys. Rev. A **1**, 18.
Allen, L. C., M. A. Paalanen, and R. N. Bhatt, 1993, Europhys. Lett. **21**, 927.
Alley, W. E., 1979, Ph.D. thesis (University of California, Davis).
Alloul, H., and P. Dellouve, 1987, Phys. Rev. Lett. **59**, 578.
Altshuler, B. L., and A. G. Aronov, 1979a, Solid State Commun. **30**, 115.
Altshuler, B. L., and A. G. Aronov, 1979b, Zh. Eksp. Teor. Fiz. **77**, 2028 [Sov. Phys. JETP **50**, 968 (1979)].
Altshuler, B. L., and A. G. Aronov, 1983, Solid State Commun. **46**, 429.
Altshuler, B. L., and A. G. Aronov, 1984, in *Electron-Electron Interactions in Disordered Systems,* edited by M. Pollak and A. L. Efros (North-Holland, Amsterdam), p. 1.
Altshuler, B. L., A. G. Aronov, D. E. Khmelnitskii, and A. I. Larkin, 1982, in *Quantum Theory of Solids,* edited by I. M. Lifshits (Mir, Moscow), p. 130.
Altshuler, B. L., A. G. Aronov, A. I. Larkin, and D. E. Khmelnitskii, 1981, Zh. Eksp. Teor. Fiz. **81**, 768 [Sov. Phys. JETP **54**, 411 (1981)].
Altshuler, B. L., A. G. Aronov, and P. A. Lee, 1980, Phys. Rev. Lett. **44**, 1288.

Altshuler, B. L., A. G. Aronov, and A. Yu. Zuzin, 1982, Pis'ma Zh. Eksp. Teor. Fiz. **35**, 15 [JETP Lett. **35**, 16 (1982)].
Altshuler, B. L., D. E. Khmelnitskii, A. I. Larkin, and P. A. Lee, 1980, Phys. Rev. B **22**, 5142.
Altshuler, B. L., V. E. Kravtsov, and I. V. Lerner, 1986, Zh. Eksp. Teor. Fiz. **91**, 2276 [Sov. Phys. JETP **64**, 1352 (1986)].
Altshuler, B. L., V. E. Kravtsov, and I. V. Lerner, 1988, Zh. Eksp. Teor. Fiz. **94**, 258 [Sov. Phys. JETP **67**, 795 (1988)].
Altshuler, B. L., V. E. Kravtsov, and I. V. Lerner, 1991, in *Mesoscopic Phenomena in Solids,* edited by B. L. Altshuler, P. A. Lee, and R. A. Webb (North-Holland, Amsterdam), p. 449.
Amit, D. J., 1984, *Field Theory, the Renormalization Group, and Critical Phenomena* (World Scientific, Singapore).
Amit, D. J., Y. Y. Goldschmidt, and L. Peliti, 1978, Ann. Phys. **116**, 1.
Anderson, P. W., 1958, Phys. Rev. **109**, 1492.
Anderson, P. W., 1959a, Phys. Rev. **115**, 2.
Anderson, P. W., 1959b, J. Phys. Chem. Solids **11**, 26.
Anderson, P. W., 1961, Phys. Rev. **124**, 41.
Anderson, P. W., 1984, *Basic Notions of Condensed Matter Physics* (Benjamin, Menlo Park).
Anderson, P. W., 1987, Science **235**, 1196.
Anderson, P. W., K. A. Muttalib, and T. V. Ramakrishnan, 1983, Phys. Rev. B **28**, 117.
Ando, T., and H. Aoki, 1985, Phys. Rev. Lett. **54**, 831.
Ando, T., A. B. Fowler, and F. Stern, 1982, Rev. Mod. Phys. **54**, 437.
Aoki, H., and T. Ando, 1985, J. Phys. Soc. Jpn. **54**, 2238.
Apel, W., 1982, J. Phys. C **15**, 1973.
Apel, W., and T. M. Rice, 1982, Phys. Rev. B **26**, 7063.
Balatsky, A. V., and E. Abrahams, 1992, Phys. Rev. B **45**, 13 125.
Balian, R., and N. R. Werthamer, 1963, Phys. Rev. **131**, 1553.
Bardeen, J., L. N. Cooper, and J. R. Schrieffer, 1957, Phys. Rev. **108**, 1175.
Béal-Monod, M. T., 1985, Phys. Rev. B **31**, 1647.
Béal-Monod, M. T., H. Ebisawa, and H. Fukuyama, 1984, Phys. Rev. B **30**, 1563.
Belitz, D., 1985, J. Phys. F **15**, 2315.
Belitz, D., 1987a, Phys. Rev. B **36**, 47.
Belitz, D., 1987b, Phys. Rev. B **36**, 2513.
Belitz, D., 1987c, Phys. Rev. B **35**, 1636.
Belitz, D., 1987d, Phys. Rev. B **35**, 1651.
Belitz, D., 1990, in *High Temperature Superconductivity,* edited by J. W. Lynn (Springer, New York), p. 29.
Belitz, D., and S. Das Sarma, 1986, Phys. Rev. B **34**, 8264.
Belitz, D., A. Gold, and W. Götze, 1981, Z. Phys. B **44**, 273.
Belitz, D., and T. R. Kirkpatrick, 1989a, Nucl. Phys. B **316**, 509.
Belitz, D., and T. R. Kirkpatrick, 1989b, Phys. Rev. Lett. **63**, 1296.
Belitz, D., and T. R. Kirkpatrick, 1991, Phys. Rev. B **44**, 955.
Belitz, D., and T. R. Kirkpatrick, 1992, Phys. Rev. B **46**, 8393.
Belitz, D., and T. R. Kirkpatrick, 1993, Phys. Rev. B **48**, 14 072.
Belitz, D., and T. R. Kirkpatrick, 1994, in "Ordering Disorder: Prospect and Retrospect in Condensed Matter Physics," edited by V. Srivastava, A. K. Bhathagar, and D. G. Naugle, AIP Conf. Proc. No. 286 (AIP, New York).
Belitz, D., and S. Q. Yang, 1993, Nucl. Phys. B **401**, 548.
Berezin, F. A., 1966, *The Method of Second Quantization* (Academic, New York).
Berezin, F. A., 1987, *Introduction to Superanalysis* (Dordrecht, Boston).
Berezinskii, V. L., 1973, Zh. Eksp. Teor. Fiz. **65**, 1251 [Sov. Phys. JETP **38**, 620 (1973)].
Berezinskii, V. L., 1974, Pis'ma Zh. Eksp. Teor. Fiz. **20**, 628 [JETP Lett. **20**, 287 (1974)].
Bergmann, G., 1969, Z. Phys. **228**, 25.
Bergmann, G., 1976, Phys. Rep. **27**, 159.
Bergmann, G., 1982, Phys. Rev. B **25**, 2937.
Bergmann, G., 1983, Phys. Rev. B **28**, 2914.
Bergmann, G., 1984, Phys. Rep. **101**, 1.
Bergmann, G., and W. Wei, 1989, Solid State Commun. **71**, 1011.
Bernas, H., and P. Nédellec, 1981, Nucl. Instrum. Methods **182/183**, 845.
Bhatt, R. N., 1993, private communication.
Bhatt, R. N., and D. S. Fisher, 1992, Phys. Rev. Lett. **68**, 3072.
Bhatt, R. N., and P. A. Lee, 1982, Phys. Rev. Lett. **48**, 344.
Biafore, M., C. Castellani, and G. Kotliar, 1990, Nucl. Phys. B **340**, 617.
Bishop, D. J., E. G. Spencer, and R. C. Dynes, 1985, Solid State Electron. **28**, 73.
Bogoliubov, N. N., V. V. Tolmachev, and D. V. Shirkov, 1959, *A New Method in the Theory of Superconductivity* (Consultants Bureau, New York).
Boon, J. P., and S. Yip, 1980, *Molecular Hydrodynamics* (McGraw-Hill, New York).
Borland, R. E. 1963, Proc. R. Soc. London, Sec. A **274**, 529.
Brézin, E., and J. Zinn-Justin, 1976, Phys. Rev. B **14**, 3110.
Brézin, E., J. Zinn-Justin, and J. C. LeGuillou, 1976, Phys. Rev. D **14**, 2615.
Broadbent, S. R., and J. M. Hammersley, 1957, Proc. Cambridge Philos. Soc. **53**, 629.
Browne, D. A., K. Levin, and K. A. Muttalib, 1987, Phys. Rev. Lett. **58**, 156.
Bruin, C., 1972, Phys. Rev. Lett. **29**, 1670.
Bruin, C., 1974, Physica (Utrecht) **72**, 261.
Bruin, C., 1978, Ph.D. thesis (Delft University).
Buckel, W., and R. Hilsch, 1954, Z. Phys. **138**, 109.
Callen, H. B., and T. A. Welton, 1951, Phys. Rev. **83**, 34.
Cardy, J. L., and H. W. Hamber, 1980, Phys. Rev. Lett. **45**, 499.
Carter, W. L., 1983, Ph.D. thesis (Stanford University).
Casher, A., D. Lurié, and M. Revzen, 1968, J. Math. Phys. **9**, 1312.
Castellani, C., 1992, private communication.
Castellani, C., and C. DiCastro, 1985, in *Localization and Metal-Insulator Transitions,* edited by H. Fritzsche and D. Adler (Plenum, New York), p. 215.
Castellani, C., and C. DiCastro, 1986, Phys. Rev. B **34**, 5935.
Castellani, C., C. DiCastro, G. Forgacs, and S. Sorella, 1984, Solid State Commun. **52**, 261.
Castellani, C., C. DiCastro, G. Kotliar, and P. A. Lee, 1986, Phys. Rev. Lett. **56**, 1179.
Castellani, C., C. DiCastro, G. Kotliar, P. A. Lee, and G. Strinati, 1987, Phys. Rev. Lett. **59**, 477.
Castellani, C., C. DiCastro, G. Kotliar, P. A. Lee, and G. Strinati, 1988, Phys. Rev. B **37**, 9046.
Castellani, C., C. DiCastro, P. A. Lee, and M. Ma, 1984, Phys. Rev. B **30**, 572.
Castellani, C., C. DiCastro, P. A. Lee, M. Ma, S. Sorella, and E. Tabet, 1984, Phys. Rev. B **30**, 1596.
Castellani, C., C. DiCastro, P. A. Lee, M. Ma, S. Sorella, and E. Tabet, 1986, Phys. Rev. B **33**, 6169.
Castellani, C., C. DiCastro, and G. Strinati, 1987, Europhys. Lett. **4**, 91.
Castellani, C., and G. Kotliar, 1986, Phys. Rev. B **34**, 9012.
Castellani, C., and G. Kotliar, 1987, Phys. Rev. B **36**, 7407.

Castellani, C., G. Kotliar, and P. A. Lee, 1987, Phys. Rev. Lett. **59**, 323.
Chakraborty, T., and P. Pietiläinen, 1988, Eds., *The Fractional Quantum Hall Effect* (Springer, New York).
Chang, K. J., M. M. Dacorogna, M. L. Cohen, and J. M. Mignot, 1985, Phys. Rev. Lett. **54**, 2375.
Chayes, J., L. Chayes, D. S. Fisher, and T. Spencer, 1986, Phys. Rev. Lett. **57**, 2999.
Chester, G. V., and A. Thellung, 1961, Proc. Phys. Soc. London **77**, 1005.
Chui, S. T., and J. W. Bray, 1977, Phys. Rev. B **16**, 1329.
Coleman, S., 1973, Commun. Math. Phys. **31**, 259.
Collins, J. C., 1984, *Renormalization* (Cambridge University, Cambridge, England).
Dai, P., Y. Zhang, S. Bogdanovich, and M. P. Sarachik, 1993, Phys. Rev. Lett. **48**, 4941.
Dai, P., Y. Zhang, and M. P. Sarachik, 1991a, Phys. Rev. Lett. **66**, 1914.
Dai, P., Y. Zhang, and M. P. Sarachik, 1991b, Phys. Rev. Lett. **67**, 136.
Dai, P., Y. Zhang, and M. P. Sarachik, 1992, Phys. Rev. B **45**, 3984.
Dai, P., Y. Zhang, and M. P. Sarachik, 1993, Phys. Rev. Lett. **70**, 1968.
Dasgupta, C., and S. K. Ma, 1980, Phys. Rev. B **22**, 1305.
Davydov, A. S., 1965, *Quantum Mechanics* (Pergamon, New York).
DeDominicis, C., and L. Peliti, 1978, Phys. Rev. B **18**, 353.
de Gennes, P. G., 1966, *Superconductivity of Metals and Alloys* (Addison-Wesley, New York).
DeGroot, S. R., and P. Mazur, 1962, *Nonequilibrium Thermodynamics* (North-Holland, Amsterdam).
De Souza, C. V., and T. R. Kirkpatrick, 1991, Phys. Rev. B **43**, 10 132.
Deutscher, G., A. M. Goldman, and H. Micklitz, 1985, Phys. Rev. B **31**, 1679.
Dobrosavljevic, V., T. R. Kirkpatrick, C. Chen, and D. Belitz, 1991, Phys. Rev. B **44**, 5432.
Dobrosavljevic, V., T. R. Kirkpatrick, and B. G. Kotliar, 1992, Phys. Rev. Lett. **69**, 1113.
Dobrosavljevic, V., and G. Kotliar, 1993, Phys. Rev. Lett. **71**, 3218.
Dolan, G. J., and D. D. Osheroff, 1979, Phys. Rev. Lett. **43**, 721.
Dorfman, J. R., and E. G. D. Cohen, 1970, Phys. Rev. Lett. **25**, 1257.
Dynes, R. C., and J. P. Garno, 1981, Phys. Rev. Lett. **46**, 137.
Dynes, R.C., A. E. White, J. M. Graybeal, and J. P. Garno, 1986, Phys. Rev. Lett. **57**, 2195.
Eckern, U., and F. Pelzer, 1988, J. Low Temp. Phys. **73**, 433.
Edwards, S. F., 1958, Philos. Mag. **3**, 1020.
Edwards, S. F., and P. W. Anderson, 1975, J. Phys. F **5**, 965.
Efetov, K. B., 1982, Zh. Eksp. Teor. Fiz. **82**, 872 [Sov. Phys. JETP **55**, 514 (1982)].
Efetov, K. B., 1983, Adv. Phys. **32**, 53.
Efetov, K. B., 1984, Zh. Eksp. Teor. Fiz. **88**, 1032 [Sov. Phys. JETP **51**, 606 (1984)].
Efetov, K. B., 1987, Zh. Eksp. Teor. Fiz. **92**, 638 [Sov. Phys. JETP **65**, 360 (1987)].
Efetov, K. B., 1988, Zh. Eksp. Teor. Fiz. **94**, 357 [Sov. Phys. JETP **67**, 199 (1988)].
Efetov, K. B., A. I. Larkin, and D. E. Khmelnitskii, 1980, Zh. Eksp. Teor. Fiz. **79**, 1120 [Sov. Phys. JETP **52**, 568 (1980)].
Efetov, K. B., and O. Viehweger, 1992, Phys. Rev. B **45**, 11 546.
Eisenriegler, E., 1973, Z. Phys. **258**, 185.
Eliashberg, G. M., 1960, Zh. Eksp. Teor. Fiz. **38**, 966 [Sov. Phys. JETP **11**, 696 (1960)].
Emery, V. J., 1979, in *Highly Conducting One-Dimensional Solids,* edited by J. T. Devreese, R. P. Evrard, and V. E. Van Doren (Plenum, New York), p. 247.
Engel, L., H.-P. Wei, D. C. Tsui, and M. Shayegan, 1990, Surface Sci. **229**, 13.
Entin-Wohlman, O., H. Gutfreund, and M. Weger, 1986, J. Phys. F **16**, 1545.
Ernst, M. H., E. H. Hauge, and J. M. J. van Leeuwen, 1970, Phys. Rev. Lett. **25**, 1254.
Ernst, M. H., and A. Weyland, 1971, Phys. Lett. A **34**, 39.
Evangelou, S. N., 1990, Physica A **167**, 199.
Evangelou, S. N., 1991, in *Quantum Coherence in Mesoscopic Systems,* edited by B. Kramer (Plenum, New York), p. 435.
Fetter, A. L., and J. D. Walecka, 1971, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, New York).
Feynman, R. P., 1949, Phys. Rev. **76**, 769.
Field, S. B., and T. F. Rosenbaum, 1985, Phys. Rev. Lett. **55**, 522.
Finkel'stein, A. M., 1983a, Zh. Eksp. Teor. Fiz. **84**, 168 [Sov. Phys. JETP **57**, 97 (1983)].
Finkel'stein, A. M., 1983b, Pis'ma Zh. Eksp. Teor. Fiz. **37**, 436 [JETP Lett. **37**, 517 (1983)].
Finkel'stein, A. M., 1984a, Zh. Eksp. Teor. Fiz. **86**, 367 [Sov. Phys. JETP **59**, 212 (1984)].
Finkel'stein, A. M., 1984b, Z. Phys. B **56**, 189.
Finkel'stein, A. M., 1984c, Pis'ma Zh. Eksp. Teor. Fiz. **40**, 63 [JETP Lett. **40**, 796 (1984)].
Finkel'stein, A. M., 1987, Pis'ma Zh. Eksp. Teor. Fiz. **45**, 37 [JETP Lett. **45**, 46 (1987)].
Finkel'stein, A. M., 1988, in *Anderson Localization,* edited by T. Ando and H. Fukuyama (Springer, New York), p. 230.
Finkel'stein, A. M., 1990, Sov. Sci. Rev. A **14**, 1.
Fisher, D. S., 1992, Phys. Rev. Lett. **69**, 534.
Fisher, D. S., M. P. A. Fisher, and D. A. Huse, 1991, Phys. Rev. B **43**, 130.
Fisher, M. E., 1983, in *Advanced Course on Critical Phenomena,* edited by F. W. Hahne (Springer, Berlin), p. 1.
Fisher, M. P. A., 1990, Phys. Rev. Lett. **65**, 923.
Fisher, M. P. A., G. Grinstein, and S. M. Girvin, 1990, Phys. Rev. Lett. **64**, 587.
Fisher, M. P. A., P. B. Weichman, G. Grinstein, and D. S. Fisher, 1989, Phys. Rev. B **40**, 546.
Forster, D., 1975, *Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions* (Benjamin, Reading, MA).
Forster, D., D. R. Nelson, and M. J. Stephen, 1977, Phys. Rev. A **16**, 732.
Friedel, J., 1956, Can. J. Phys. **34**, 1190.
Fröhlich, J., F. Martinelli, E. Scoppola, and T. Spencer, 1985, Commun. Math. Phys. **101**, 21.
Fukuyama, H., 1980, J. Phys. Soc. Jpn. **48**, 2169.
Fukuyama, H., H. Ebisawa, and S. Maekawa, 1984, J. Phys. Soc. Jpn. **53**, 3560.
Furubayashi, T., N. Nishida, M. Yamaguchi, K. Morigaki, and H. Ishimoto, 1985, Solid State Commun. **55**, 513.
Fyodorov, Y. V., and A. D. Mirlin, 1991, Phys. Rev. Lett. **67**, 2049.
Gan, Z.-Z., and P. A. Lee, 1986, Phys. Rev. B **33**, 3595.
Gell-Mann, M., and M. Lévy, 1960, Nuovo Cimento **16**, 1729.
Georges, A., and W. Krauth, 1992, Phys. Rev. Lett. **69**, 1240.
Giamarchi, T., and H. J. Schulz, 1988, Phys. Rev. B **37**, 325.
Giuliani, G. F., and J. J. Quinn, 1984, Phys. Rev. B **29**, 2321.

Glauber, R., 1963, Phys. Rev. **131**, 2766.
Gold, A., 1984, Ph.D. thesis (Technical University Munich).
Gold, A., 1985a, Phys. Rev. Lett. **54**, 1079.
Gold, A., 1985b, Phys. Rev. B **32**, 4014.
Gold, A., and W. Götze, 1983a, Helv. Phys. Acta **56**, 47.
Gold, A., and W. Götze, 1983b, Solid State Commun. **47**, 627.
Gold, A., and W. Götze, 1985, Solid State Electron. **28**, 87.
Gold, A., and W. Götze, 1986, Phys. Rev. B **33**, 2495.
Gold, A., W. Götze, C. Mazuré, and F. Koch, 1984, Solid State Commun. **49**, 1085.
Gorkov, L. P., 1959, Zh. Eksp. Teor. Fiz. **37**, 1407 [Sov. Phys. JETP **37**, 998 (1959)].
Gorkov, L. P., A. Larkin, and D. E. Khmelnitskii, 1979, Pis'ma Zh. Eksp. Teor. Fiz. **30**, 248 [JETP Lett. **30**, 228 (1979)].
Götze, W., 1981, Philos. Mag. B **43**, 219.
Götze, W., and W. Ketterle, 1983, Z. Phys. B **54**, 49.
Götze, W., E. Leutheusser, and S. Yip, 1982, Phys. Rev. A **25**, 533.
Graybeal, J. M., 1985, Ph.D. thesis (Stanford University).
Graybeal, J. M., and M. R. Beasley, 1984, Phys. Rev. B **29**, 4167.
Graybeal, J. M., M. R. Beasley, and R. L. Greene, 1984, in *Proceedings of the LT17,* edited by U. Eckern, A. Schmid, and W. Weber (North-Holland, Amsterdam), p. 731.
Grilli, M., and S. Sorella, 1988, Nucl. Phys. B **295** [FS 21], 422.
Grinstein, G., 1985, in *Fundamental Problems in Statistical Mechanics VI,* edited by E. G. D. Cohen (North-Holland, Amsterdam), p. 147.
Grüner, G., and A. Zawadowski, 1974, Rep. Prog. Phys. **37**, 1497.
Grünewald, G., and K. Scharnberg, 1974, Z. Phys. **268**, 197.
Grünewald, G., and K. Scharnberg, 1975, Z. Phys. B **20**, 61.
Gutfreund, H., M. Weger, and O. Entin-Wohlman, 1985, Phys. Rev. B **31**, 606.
Gutzwiller, M., 1963, Phys. Rev. Lett. **10**, 159.
Haldane, F. D. M., 1978, Phys. Rev. Lett. **40**, 489.
Hamermesh, M., 1962, *Group Theory* (Addison-Wesley, Reading, MA).
Harris, A. B., and T. C. Lubensky, 1981, Phys. Rev. B **23**, 2640.
Hauge, E. H., 1974, in *Transport Phenomena,* Lecture Notes in Physics No. 31, edited by G. Kirczenow and J. Marro (Springer, New York), p. 337.
Haviland, D. B., Y. Liu, and A. M. Goldman, 1989, Phys. Rev. Lett. **62**, 2180.
Hebard, A. F., and G. Kotliar, 1989, Phys. Rev. B **39**, 4105.
Hebard, A. F., and M. A. Paalanen, 1985, Phys. Rev. Lett. **54**, 2155.
Hebard, A. F., and M. A. Paalanen, 1990, Phys. Rev. Lett. **65**, 927.
Hebard, A. F., and M. A. Paalanen, 1992, Helv. Phys. Acta **65**, 197.
Hertz, J., 1976, Phys. Rev. B **14**, 1165.
Hikami, S., 1980, Prog. Theor. Phys. **64**, 1466.
Hikami, S., 1982, in *Anderson Localization,* edited by Y. Nagaoka and H. Fukuyama (Springer, New York), p. 15.
Hikami, S., 1990, Physica A **167**, 149.
Hikami, S., 1992, Prog. Theor. Phys. Suppl. **107**, 213.
Hikami, S., and E. Brézin, 1978, J. Phys. A **11**, 1141.
Hikami, S., A. I. Larkin, and Y. Nagaoka, 1980, Prog. Theor. Phys. **63**, 707.
Hirsch, M. J., D. F. Holcomb, R. N. Bhatt, and M. A. Paalanen, 1992, Phys. Rev. Lett. **68**, 1418.
Hirsch, M. J., U. Thomanschefsky, and D. F. Holcomb, 1988, Phys. Rev. B **37**, 8257.
Hoch, M. J. R., and D. F. Holcomb, 1988, Phys. Rev. B **38**, 10 550.
Höf, D., and F. Wegner, 1986, Nucl. Phys. B **275** [FS 17], 561.
Hohenberg, P. C., 1967, Phys. Rev. **158**, 383.
Hohenberg, P. C., and B. I. Halperin, 1977, Rev. Mod. Phys. **49**, 435.
Holstein, T., 1959, Phys. Rev. **113**, 479.
Houghton, A., L. Schäfer, and F. Wegner, 1980, Phys. Rev. B **22**, 3598.
Houghton, A., and H. Won, 1985, J. Phys. C **18**, 2507.
Hubbard, J., 1959, Phys. Rev. Lett. **3**, 77.
Hubbard, J., 1963, Proc. R. Soc. London, Sec. A **276**, 238.
Huckestein, B., and B. Kramer, 1990, Phys. Rev. Lett. **64**, 1437.
Iyzykson, C., and J.-M. Drouffe, 1989, *Statistical Field Theory* (Cambridge University, Cambridge, England).
Jaeger, H. M., D. B. Haviland, A. M. Goldman, and B. G. Orr, 1986, Phys. Rev. B **34**, 4920.
Jain, J. K., 1992, Adv. Phys. **141**, 105.
Jarrel, M., 1992, Phys. Rev. Lett. **69**, 168.
Jerome, D., C. Ryter, H. J. Schulz, and J. Friedel, 1985, Philos. Mag. B **52**, 403.
Jisrawi, N., W. L. McLean, and N. G. Stoffel, 1987, in *Novel Superconductivity,* edited by S. A. Wolf and V. Z. Kresin (Plenum, New York), p. 29.
Kadanoff, L. P., and G. Baym, 1962, *Quantum Statistical Mechanics* (Benjamin, New York).
Kadanoff, L. P., and I. I. Falko, 1964, Phys. Rev. **136**, 1170A.
Kapitulnik, A., and G. Kotliar, 1985, Phys. Rev. Lett. **54**, 473.
Katsumoto, S., F. Komori, N. Sano, and S. Kobayashi, 1987, J. Phys. Soc. Jpn. **56**, 2259.
Katsumoto, S., F. Komori, N. Sano, and S. Kobayashi, 1989, Solid State Commun. **71**, 441.
Kaveh, M., and N. F. Mott, 1987, Philos. Mag. B **56**, 97.
Kawabata, A., 1981, Solid State Commun. **38**, 823.
Keck, B., and A. Schmid, 1975, Solid State Commun. **17**, 799.
Keck, B., and A. Schmid, 1976, J. Low. Temp. Phys. **24**, 611.
Khinchin, A. J., 1949, *Mathematical Foundations of Statistical Mechanics* (Dover, New York).
Kim, K., and P. B. Weichmann, 1991, Phys. Rev. B **43**, 13 583.
Kirkpatrick, T. R., and D. Belitz, 1986a, Phys. Rev. B **34**, 2168.
Kirkpatrick, T. R., and D. Belitz, 1986b, Phys. Rev. B **34**, 9008.
Kirkpatrick, T. R., and D. Belitz, 1989, Phys. Rev. B **40**, 5227.
Kirkpatrick, T. R., and D. Belitz, 1990a, Phys. Rev. B **41**, 11 082.
Kirkpatrick, T. R., and D. Belitz, 1990b, J. Phys. Condens. Matter **2**, 5259.
Kirkpatrick, T. R., and D. Belitz, 1991, Phys. Rev. Lett. **66**, 1533.
Kirkpatrick, T. R., and D. Belitz, 1992a, J. Phys. Condens. Matter. **4**, L37.
Kirkpatrick, T. R., and D. Belitz, 1992b, Phys. Rev. B **45**, 3187.
Kirkpatrick, T. R., and D. Belitz, 1992c, Phys. Rev. Lett. **68**, 3232.
Kirkpatrick, T. R., and D. Belitz, 1993, Phys. Rev. Lett. **70**, 974.
Kirkpatrick, T. R., and D. Belitz, 1994, "Cooperons at the metal-insulator transition revisited: Constraints on the renormalization group, and a conjecture," University of Oregon preprint.
Kirkpatrick, T. R., and J. R. Dorfman, 1983, Phys. Rev. A **22**, 1022.
Koba, Z., N. Mugibayashi, and S. Nakai, 1951, Prog. Theor. Phys. **6**, 322.
Koch, S., R. G. Haug, K. von Klitzing, and K. Ploog, 1992,

Phys. Rev. B **46**, 1596.
Kondo, J., 1964, Prog. Theor. Phys. **32**, 37.
Koon, D., and T. Castner, 1988, Phys. Rev. Lett. **60**, 1755.
Korringa, J., 1950, Physica **16**, 601.
Kotliar, G., 1992, private communication.
Kotliar, G., and A. Kapitulnik, 1986, Phys. Rev. B **33**, 3146.
Kotliar, B. G., and T. V. Ramakrishnan, 1985, Phys. Rev. B **31**, 8188.
Kramer, B., K. Broderix, A. MacKinnon, and M. Schreiber, 1990, Physica A **167**, 163.
Kravtsov, V. E., I. V. Lerner, and V. I. Yudson, 1988, Zh. Eksp. Teor. Fiz. **94**, 255 [Sov. Phys. JETP **68**, 1441 (1988)].
Kravtsov, V. E., and R. Oppermann, 1991, Phys. Rev. B **43**, 10 865.
Krishnamurthy, H. R., J. W. Wilkins, and K. G. Wilson, 1980a, Phys. Rev. B **21**, 1003.
Krishnamurthy, H. R., J. W. Wilkins, and K. G. Wilson, 1980b, Phys. Rev. B **21**, 1044.
Kubo, R., 1957, J. Phys. Soc. Jpn. **12**, 570.
Lakner, M., and H. von Löhneysen, 1989, Phys. Rev. Lett. **63**, 648.
LaMadrid, M. A., W. Contrata, and J. M. Mochel, 1992, Phys. Rev. B **45**, 3870.
Landau, L. D., 1958, Zh. Eksp. Teor. Fiz. **35**, 97 [Sov. Phys. JETP **35**, 70 (1958)].
Langer, J. S., and T. Neal, 1966, Phys. Rev. Lett. **16**, 984.
Leavens, C. R., 1985, Phys. Rev. B **31**, 6072.
Le Bellac, M., 1991, *Quantum and Statistical Field Theory* (Clarendon, Oxford).
Lee, P. A., 1982, Phys. Rev. B **26**, 5882.
Lee, P. A., and T. V. Ramakrishnan, 1985, Rev. Mod. Phys. **57**, 287.
Lee, P. A., A. D. Stone, and H. Fukuyama, 1987, Phys. Rev. B **35**, 1039.
Lee, S. J., and J. B. Ketterson, 1990, Phys. Rev. Lett. **64**, 3078.
Lepage, G. P., 1990, in *From Actions to Answers,* edited by T. DeGrand and D. Toussaint (World Scientific, Singapore), p. 483.
Lerner, I. V., 1988, Phys. Lett. A **133**, 253.
Lerner, I. V., and F. Wegner, 1990, Z. Phys. B **81**, 95.
Lesueur, J., L. Dumoulin, and P. Nédellec, 1988, Solid State Commun. **66**, 723.
Levelt Sengers, J. M. H., and J. V. Sengers, 1981, in *Perspectives in Statistical Physics,* edited by H. J. Raveche (North-Holland, Amsterdam), p. 241.
Levin, K., and O. T. Valls, 1978, Phys. Rev. B **17**, 191.
Lieb, E. H., 1989, Phys. Rev. Lett. **62**, 1201.
Lieb, E. H., and F. Y. Wu, 1968, Phys. Rev. Lett. **20**, 1445.
Lindhard, J., 1954, K. Dan. Vidensk. Selsk.-Mat.-Fys. Medd. **28**, Nr. 8, p. 1.
Liu, D., and S. Das Sarma, 1993, Mod. Phys. Lett. B **7**, 449.
Long, A. P., and M. Pepper, 1984, J. Phys. C **17**, L425.
Lynton, E. A., B. Serin, and M. Zucker, 1957, J. Phys. Chem. Solids **3**, 165.
Ma, M., and E. Fradkin, 1986, Phys. Rev. Lett. **56**, 1416.
Ma, M., B. I. Halperin, and P. A. Lee, 1986, Phys. Rev. B **34**, 3136.
Ma, M., and P. A. Lee, 1985, Phys. Rev. B **32**, 5658.
Ma, S.-K., 1976, *Modern Theory of Critical Phenomena* (Benjamin, Reading, MA).
MacKinnon, A., and B. Kramer, 1993, Rep. Prog. Phys. **56**, 1469.
Maekawa, S., H. Ebisawa, and H. Fukuyama, 1984, J. Phys. Soc. Jpn. **53**, 2681.
Maekawa, S., and H. Fukuyama, 1981, Physica B **107**, 123.
Maekawa, S., and H. Fukuyama, 1982, J. Phys. Soc. Jpn. **51**, 1380.
Mahan, G. D., 1981, *Many-Particle Physics* (Plenum, New York).
Matthias, B. T., H. Suhl, and E. Corenzwit, 1958, Phys. Rev. Lett. **1**, 92.
Matthias, B. T., E. A. Wood, E. Corenzwit, and V. B. Bala, 1956, J. Phys. Chem. Solids **1**, 188.
McKane, A. J., and M. Stone, 1981, Ann. Phys. (N.Y.) **131**, 36.
McMillan, W. L., 1981, Phys. Rev. B **24**, 2739.
McMillan, W. L., and J. Mochel, 1981, Phys. Rev. Lett. **46**, 556.
Mermin, N. D., and H. Wagner, 1966, Phys. Rev. Lett. **17**, 1133.
Metzner, W., and D. Vollhardt, 1989, Phys. Rev. Lett. **62**, 324.
Miehle, W., R. Gerber, and P. Ziemann, 1992, Phys. Rev. B **45**, 895.
Milovanovich, M., S. Sachdev, and R. N. Bhatt, 1989, Phys. Rev. Lett. **63**, 82.
Morel, P., and P. W. Anderson, 1962, Phys. Rev. **125**, 1263.
Mott, N. F., 1949, Proc. Phys. Soc. London, Ser. A **62**, 416.
Mott, N. F., 1966, Philos. Mag. **13**, 989.
Mott, N. F., 1990, *Metal-Insulator Transitions* (Taylor & Francis, London).
Mott, N. F., and W. D. Twose, 1961, Adv. Phys. **10**, 107.
Negele, J. W., and H. Orland, 1988, *Quantum Many-Particle Systems* (Addison-Wesley, New York).
Nelson, D. R., and R. A. Pelcovits, 1977, Phys. Rev. B **16**, 2191.
Nelson, D. R., and J. Rudnick, 1976, Phys. Rev. B **13**, 2208.
Newman, P. F., and D. F. Holcomb, 1983a, Phys. Rev. B **28**, 638.
Newman, P. F., and D. F. Holcomb, 1983b, Phys. Rev. Lett. **51**, 2144.
Nickerson, H. K., D. C. Spencer, and N. E. Steenrod, 1959, *Advanced Calculus* (Van Nostrand, Princeton).
Nishida, N., T. Furubayashi, M. Yamaguchi, K. Morigaki, and H. Ishimoto, 1985, Solid State Electron. **28**, 81.
Nishida, N., T. Furubayashi, M. Yamaguchi, K. Morigaki, Y. Miura, Y. Takano, H. Ishimoto, and S. Ogawa, 1984, in *Proceedings of the LT-17,* edited by U. Eckern, A. Schmid, W. Weber, and H. Wühl (Elsevier, Amsterdam), p. 731.
Nozières, P., 1979, J. Phys. C **39**, 117.
Ootuka, Y., and N. Matsunaga, 1990, J. Phys. Soc. Jpn. **59**, 1801.
Oppermann, R., 1985, Z. Phys. B **61**, 89.
Oppermann, R., 1986, Z. Phys. B **63**, 33.
Oppermann, R., 1987a, Europhys. Lett. **3**, 431.
Oppermann, R., 1987b, Nucl. Phys. B **280**, 753.
Oppermann, R., 1987c, Z. Phys. B **68**, 329.
Oppermann, R., 1988a, Z. Phys. B **70**, 43.
Oppermann, R., 1988b, Z. Phys. B **70**, 49.
Oppermann, R., 1988c, Solid State Commun. **65**, 1391.
Oppermann, R., 1990, Physica A **167**, 301.
Ovchinnikov, Yu. N., 1973, Zh. Eksp. Teor. Fiz. **64**, 719 [Sov. Phys. JETP **37**, 366 (1973)].
Paalanen, M. A., J. E. Graebner, R. N. Bhatt, and S. Sachdev, 1988, Phys. Rev. Lett. **61**, 597.
Paalanen, M. A., T. F. Rosenbaum, G. A. Thomas, and R. N. Bhatt, 1982, Phys. Rev. Lett. **48**, 1284.
Paalanen, M. A., A. E. Ruckenstein, and G. A. Thomas, 1985, Phys. Rev. Lett. **54**, 1295.
Paalanen, M. A., S. Sachdev, R. N. Bhatt, and A. E. Ruckenstein, 1986, Phys. Rev. Lett. **57**, 2061.
Pao, C.-H., and N. E. Bickers, 1992, Phys. Rev. B **44**, 10 270.

Parks, R. D., 1969, Ed., *Superconductivity,* Vols. I and II (Dekker, New York).
Pelcovits, R. A., and D. R. Nelson, 1976, Phys. Lett. A **57**, 23.
Pines, D., and P. Nozières, 1989, *The Theory of Quantum Liquids* (Addison-Wesley, Reading, MA).
Pippard, A. B., 1955, Philos. Mag. **46**, 1104.
Polyakov, A. M., 1975, Phys. Lett. B **59**, 79.
Pomeau, Y., and P. Résibois, 1975, Phys. Rep. **19**, 63.
Popov, V. N., 1983, *Functional Integrals in Quantum Field Theory and Statistical Physics* (Reidel, Dordrecht).
Prange, R. E., and S. M. Girvin, 1987, Eds., *The Quantum Hall Effect* (Springer, New York).
Press, N. H. B. P. Flannery, S. A. Teutolsky, and W. T. Vetterling, 1986, *Numerical Recipes* (Cambridge University, Cambridge, England).
Pruisken, A. M. M., 1984, Nucl. Phys. B **235**, 277.
Pruisken, A. M. M., 1988, Phys. Rev. Lett. **61**, 1297.
Pruisken, A. M. M., and L. Schäfer, 1982, Nucl. Phys. B **200** [FS 4], 20.
Pruisken, A. M. M., and Z. Wang, 1989, Nucl. Phys. B **322**, 721.
Quirt, J. D., and J. R. Marko, 1971, Phys. Rev. Lett. **26**, 318.
Raimondi, R., C. Castellani, and C. DiCastro, 1990, Phys. Rev. B **42**, 4724.
Ramakrishnan, T. V., 1989, Phys. Scr. **T27**, 24.
Reizer, M. Yu., and A. V. Sergeyev, 1986, Zh. Eksp. Teor. Fiz. **90**, 1056 [Sov. Phys. JETP **63**, 616 (1986)].
Rhode, M., and H. Micklitz, 1987, Phys. Rev. B **36**, 7572.
Rosenbaum, T. F., K. Andres, G. A. Thomas, and P. A. Lee, 1981, Phys. Rev. Lett. **46**, 568.
Rosenbaum, T. F., R. F. Milligan, M. A. Paalanen, G. A. Thomas, R. N. Bhatt, and W. Lin, 1983, Phys. Rev. B **27**, 7509.
Rosenbaum, T. F. G. A. Thomas, and M. A. Paalanen, 1994, Phys. Rev. Lett. **72**, 2121.
Rowell, J. M., and R. C. Dynes, 1980, *Bad Metals, Good Superconductors* (unpublished).
Rozenberg, M. J., X. Y. Zhang, and G. Kotliar, 1992, Phys. Rev. Lett. **69**, 1236.
Sachdev, S., 1989, Phys. Rev. B **39**, 5297.
Sarachik, M. P., 1992, private communication.
Sarachik, M. P., D. R. He, W. Li, M. Levy, and J. S. Brooks, 1985, Phys. Rev. B **31**, 1469.
Scalapino, D. J., J. R. Schrieffer, and J. W. Wilkins, 1966, Phys. Rev. **148**, 263.
Schäfer, L., and F. Wegner, 1980, Z. Phys. B **38**, 113.
Schmid, A., 1970, Z. Phys. **231**, 324.
Schmid, A., 1973, Z. Phys. **259**, 421.
Schmid, A., 1974, Z. Phys. **271**, 251.
Schmid, A., 1985, in "Localization, Interaction, and Transport Phenomena," edited by B. Kramer, G. Bergmann, and Y. Bruynseraede (Springer, Berlin), p. 212.
Schönhammer, K., and W. Brenig, 1973, Phys. Lett. A **42**, 447.
Schrieffer, J. R., 1964, *Superconductivity* (Benjamin, Reading, MA).
Shafarman, W. N., D. W. Koon, and T. G. Castner, 1989, Phys. Rev. B **40**, 1216.
Si, Q., and G. Kotliar, 1993, Phys. Rev. Lett. **70**, 3143.
Smith, R. A., and V. Ambegaokar, 1992, Phys. Rev. B **45**, 2463.
Sørensen, E. S., M. Wallin, S. M. Girvin, and A. P. Young, 1992, Phys. Rev. Lett. **69**, 828.
Sota, T., and K. Suzuki, 1989, Phys Rev. B **40**, 2464.
Stratonovich, R. L., 1957, Dok. Akad. Nauk. SSSR **115**, 1907 [Sov. Phys. Dokl. **2**, 416 (1957)].
Strinati, G., and C. Castellani, 1987, Phys. Rev. B **36**, 2270.
Strongin, M., R. S. Thompson, O. F. Kammerer, and J. E. Crow, 1970, Phys. Rev. B **1**, 1078.
Stupp, H., M. Hornung, M. Lakner, O. Madel, and H. v. Löhneysen, 1994a, Phys. Rev. Lett. **71**, 2634.
Stupp, H., M. Hornung, M. Lakner, O. Madel, and H. v. Löhneysen, 1994b, Phys. Rev. Lett. **72**, 2122.
Suzumura, Y., and H. Fukuyama, 1983, J. Phys. Soc. Jpn. **52**, 2870.
Symanzik, K., 1983, Nucl. Phys. B **226**, 187.
Takagi, H., and Y. Kuroda, 1982, Solid State Commun. **41**, 643.
Testardi, L. R., and L. F. Mattheiss, 1978, Phys. Rev. Lett. **41**, 1612.
Thomas, G. A., 1985, Philos. Mag. B **52**, 479.
Thomas, G. A., M. Capizzi, F. DeRosa, R. N. Bhatt, and T. M. Rice, 1981, Phys. Rev. B **23**, 5472.
Thomas, G. A., M. A. Paalanen, and T. F. Rosenbaum, 1983, Phys. Rev. B **27**, 3897.
Tinkham, M., 1975, *Introduction to Superconductivity* (McGraw-Hill, New York).
Tsuneto, T., 1961, Phys. Rev. **121**, 402.
Tunstall, D. P. 1984, J. Phys. C **17**, L1003.
Tunstall, D. P., and V. G. I. Deshmukh, 1979, J. Phys. C **12**, 2295.
Tunstall, D. P., and G. S. Sohal, 1983, J. Phys. C **16**, L251.
Ue, H., and S. Maekawa, 1971, Phys. Rev. B **3**, 4232.
Valles, J. M., Jr., 1992, private communication.
Valles, J. M., Jr., and R. C. Dynes, 1990, in *Physical Phenomena in Granular Materials,* MRS Symposium Series, No. 195, edited by G. D. Cody, T. H. Geballe, and Ping Sheng (Materials Research Society, Pittsburgh, PA), p. 375.
Valles, J. M., Jr., R. C. Dynes, and J. P. Garno, 1989, Phys. Rev. B **40**, 6680.
van der Waerden, B. L., 1970, *Algebra* (Frederick Ungar, New York).
Van Leeuwen, J. M. J., and A. Weyland, 1987, Physics A **36**, 457.
Verbaarschot, J. J. M., and M. R. Zirnbauer, 1985, J. Phys. A **18**, 1093.
Vollhardt, D., and P. Wölfle, 1980, Phys. Rev. B **22**, 4666.
Vollhardt, D., and P. Wölfle, 1982, Phys. Rev. Lett. **48**, 699.
Vollhardt, D., and P. Wölfle, 1990, *The Superfluid Phases of Helium 3* (Taylor & Francis, London).
von Klitzing, K., G. Dorda, and M. Pepper, 1980, Phys. Rev. Lett. **45**, 494.
Wang, X.-F., Z. Wang, G. Kotliar, and C. Castellani, 1992, Phys. Rev. Lett. **68**, 2504.
Wegner, F., 1972, Phys. Rev. B **5**, 4529.
Wegner, F., 1976a, Z. Phys. B **25**, 327.
Wegner, F., 1976b, in *Phase Transitions and Critical Phenomena,* edited by C. Domb and M. S. Green (Academic, New York), Vol. 6, p. 7.
Wegner, F., 1979, Z. Phys. B **35**, 207.
Wegner, F., 1980, Z. Phys. B **36**, 209.
Wegner, F., 1981a, Z. Phys. B **44**, 9.
Wegner, F., 1981b, Nucl. Phys. B **180** [FS 2], 77.
Wegner, F., 1986, Nucl. Phys. B **270** [FS 16], 1.
Wegner, F., 1987a, Nucl. Phys. B **280** [FS 18], 193.
Wegner, F., 1987b, Nucl. Phys. B **280** [FS 18], 210.
Wegner, F., 1989, Nucl. Phys. B **316**, 663.
Wegner, F., 1990, Z. Phys. B **78**, 33.
Wei, H. P., D. C. Tsui, M. A. Paalanen, and A. M. M. Pruisken, 1988, Phys. Rev. Lett. **61**, 1294.
Weinberg, S., 1983, in *Asymptotic Realms of Physics,* edited by A. H. Guth, K. Huang, and R. L. Jaffe (MIT, Cambridge, MA), p. 1.

Wilson, K. G., and J. Kogut, 1974, Phys. Rep. **12**, 75.
Wolff, P. A., 1961, Phys. Rev. **124**, 1030.
Wölfle, P., and D. Vollhardt, 1982, in *Anderson Localization,* edited by Y. Nagaoka and H. Fukuyama (Springer, New York), p. 26.
Wysokinski, K. I., and A. L. Kuzemski, 1983, J. Low. Temp. Phys. **52**, 81.
Wysokinski, K. I., W. Park, D. Belitz, and T. R. Kirkpatrick, 1994, unpublished.
Yoshizumi, S., D. Mael, T. H. Geballe, and R. L. Greene, 1985, in *Localization and Metal-Insulator Transitions,* edited by H. Fritzsche and D. Adler (Plenum, New York), p. 77.
Ziman, J. M., 1964, *Principles of the Theory of Solids* (Cambridge University, Cambridge, England).
Zimanyi, G. T., and E. Abrahams, 1990, Phys. Rev. Lett. **64**, 2719.
Zinn-Justin, J., 1989, *Quantum Field Theory and Critical Phenomena* (Clarendon, Oxford).
Zint, T., M. Rhode, and M. Micklitz, 1990, Phys. Rev. B **41**, 4831.
Zirnbauer, M. R., 1986, Phys. Rev. B **34**, 6394.
Zirnbauer, M. R., 1988, Phys. Rev. Lett. **60**, 1450.
Zubarev, D. N., 1960, Usp. Fiz. Nauk. **71**, 71 [Sov. Phys. Usp. **3**, 320 (1960)].