Nonanalyticities in a strongly correlated Fermi liquid: Corrections to scaling at the Fermi-liquid fixed point

D. Belitz[1,2] and T. R. Kirkpatrick[3]

[1]*Department of Physics and Institute of Theoretical Science, University of Oregon, Eugene, Oregon 97403, USA*

[2]*Materials Science Institute, University of Oregon, Eugene, Oregon 97403, USA*

[3]*Institute for Physical Science and Technology and Department of Physics, University of Maryland, College Park, Maryland 20742, USA*

(Received 27 August 2013; revised manuscript received 28 December 2013; published 21 January 2014)

We use scaling and renormalization-group techniques to analyze the leading nonanalyticities in a Fermi liquid. We show that a physically motivated scaling hypothesis reproduces the results known from perturbation theory for the density of states, the density-of-states fluctuations, the specific heat, the spin susceptibility, and the nematic magnetic susceptibility. We also discuss the absence of nonanalytic terms in the density susceptibility. We then use a recent effective field theory for clean electron systems to derive the scaling hypothesis by means of renormalization-group techniques. This shows that the exponents (although not the prefactors) of the nonanalyticities that were previously derived by means of perturbative techniques are indeed exact, and can be understood as the leading corrections to scaling at the stable Fermi-liquid fixed point.

DOI: 10.1103/PhysRevB.89.035130 PACS number(s): 71.27.+a, 05.30.Fk, 71.10.−w, 71.45.−d

I. INTRODUCTION AND RESULTS

A. Introduction

In any system with soft or massless excitations, the observable behavior at long wavelengths and low frequencies is dominated by these soft modes, to the extent that the observables couple to them. In the quantum regime, the soft modes in addition govern the low-temperature behavior. In some systems, soft modes exist only at special points in the phase diagram, for instance, at the critical point of a second-order phase transition. An example are Ising magnets. In others, soft modes exist in entire phases, either because a spontaneously broken continuous symmetry leads to Goldstone modes or because of conservation laws. We will refer to such soft modes as "generic." Examples are the ferromagnons in the ordered phase of a Heisenberg ferromagnet. Technically, scaling ideas and the renormalization group (RG) are well suited to deal with soft modes and their consequences. In condensed matter physics, RG techniques are best known for the theory of singularities at critical points [1]. However, it arguably is even more interesting and important to understand singularities that appear in entire phases. The fact that the RG is equally useful for this purpose is less well known. In the RG framework, critical singularities are described in terms of critical fixed points of the RG transformations, whereas entire phases are described in terms of stable fixed points [2,3]. At the former, the control parameter (in the case of a thermal phase transition, the temperature deviation from the critical temperature) is the only relevant operator; at the latter, there are no relevant operators (ignoring external fields in both cases).

There are many examples of long-wavelength and/or low-frequency singularities that exist in entire phases due to the coupling of observables to generic soft modes [4], and we list only a few for illustration purposes: (1) The longitudinal magnetic susceptibility χ_L in the ordered phase of an isotropic Heisenberg ferromagnet in $d < 4$ dimensions diverges as $\chi_L(h \to 0) \propto h^{-(4-d)/2}$ as a result of the longitudinal magnetization coupling to the magnons [5]. (2) The tunneling density of states (DOS) N in a disordered electron system as a function of the energy distance (or bias voltage) ω from the Fermi energy ϵ_F shows a cusp singularity $N(\epsilon_F + \omega) \propto \text{const} + |\omega|^{(d-2)/2}$ as a result of a coupling to the diffusive modes known as "diffusons"[6,7]. At the Fermi energy (zero bias), the tunneling DOS shows the same behavior as a function of the temperature T, with $|\omega|$ replaced by T [8]. In $d \leqslant 2$ the singularity is so strong that the disordered Fermi liquid is destroyed. (3) The kinematic viscosity ν in a classical fluid is a nonanalytic function of the frequency ω: $\nu(\omega \to 0) \propto \text{const} + \omega^{(d-2)/2}$. This "long-time tail" (the corresponding time correlation function decays as $1/t^{d/2}$ for long times t, rather than exponentially) [9–12] is a consequence of the coupling to various soft modes in fluids that exist as a result of conservation laws. Again, in $d \leqslant 2$ the strong singularity leads to a breakdown of local hydrodynamics. These are just three examples each of three rather large classes of phenomena in classical magnets, disordered electrons, and classical fluids, respectively, all of which can be studied by RG techniques (see Appendices A, B, and C).

These analogies between classical and quantum systems notwithstanding, there also are important qualitative differences between them. In classical systems, the statics and the dynamics are decoupled in equilibrium, meaning that equilibrium static correlation functions can be long ranged only due Goldstone modes. Only in a nonequilibrium situation do the statics and dynamics couple, and long-range static correlations can result from conservation laws as well [13]. In quantum systems, on the other hand, the statics and dynamics are intrinsically coupled even in equilibrium, and long-range static correlations can result from either Goldstone modes or conservation laws.

In this paper, we use RG and scaling techniques to discuss the interesting singularities that occur in a clean Fermi liquid in dimensions $d > 1$. It is well known that the leading behavior of interacting electrons at low temperature, frequency, and wave number is described by Landau's Fermi-liquid theory [14], which corresponds to a stable fixed point in a RG framework [15]. One major point of this paper is to show that various nonanalytic corrections to Fermi-liquid

theory that have been previously derived using perturbation theory can be understood as the leading corrections to scaling at this stable fixed point. As we will see, there are strong technical and physical analogies between these singularities and the ones in disordered Fermi liquids, and to a lesser extent in classical magnets and classical fluids, mentioned above. Another important point is that RG techniques will allow us to show that the perturbative results are actually exact as far as the exponents that characterize the singularities are concerned. In $d = 1$ the singularities we will discuss are so strong that they contribute to an instability of the Fermi liquid in favor of a Luttinger-liquid state [16]. In strongly correlated systems, they may lead to a quantum phase transition from a Fermi liquid to a non-Fermi-liquid state even in $d > 1$ [17].

The RG methods we employ are the "traditional" ones as described in Refs. [1,2]. We specify the order parameter, identify all of the soft modes that couple to it, and then focus on a description of these modes and the order-parameter fluctuations. This is in contrast to the functional RG methods that in recent years have been applied to fermion systems (see Ref. [18] and references therein). The latter approach uses an unbiased RG method that does not *a priori* single out a particular order parameter and does not focus on soft modes. It also does not involve a rescaling procedure, or introduce the notion of relevant versus irrelevant variables. Rather, it starts from a microscopic Hamiltonian or action and examines the scale dependencies of all quantities, universal or nonuniversal, using a variety of truncation schemes. It is well suited to study situations where various instabilities compete with one another. In contrast, the RG approach used in this paper is much more restrictive, but also much easier to use, as it focuses on the universal long-wavelength properties of the system. Specifically, we explicitly identify the Goldstone modes that characterize the Fermi-liquid phase, and we show that these modes are the physical origin of the leading corrections to scaling at the stable Fermi-liquid fixed point. The stability of these modes, and some properties of a quantum phase transition from a Fermi liquid to a non-Fermi-liquid phase, have been considered elsewhere [17].

This paper is organized as follows. In Sec. II we formulate a scaling hypothesis for the free energy. From this we derive homogeneity laws for the observables of interest, which all can be expressed as derivatives of the free energy. The resulting leading nonanalytic corrections to Fermi-liquid theory are all consistent with previous results from explicit perturbative calculations. In Sec. III we recall the schematic form of a recent effective field theory for clean fermions [19]. We identify the fixed-point action that describes the Fermi liquid, and identify the leading irrelevant operators with respect to this fixed point. This allows for a derivation of the scaling hypothesis, and hence shows that the exponents derived from the scaling considerations are exact. In Sec. IV we discuss various aspects of our approach, and our results. Analogies with classical magnets, disordered electron systems, and classical fluids, respectively, that are of physical or pedagogical interest are the subject of Appendices A, B, and C. Appendix D contains technical details regarding the density of states and its susceptibility in the presence of a long-range Coulomb interaction, and Appendix E explains a structural feature of the density-of-states susceptibility.

II. PHENOMENOLOGICAL SCALING CONSIDERATIONS FOR THE FREE ENERGY AND ITS DERIVATIVES

Before we go into the technical details of applying RG ideas to analyze a microscopic theory, let us employ simple phenomenological scaling arguments [20] to find out what one should expect for the behavior of various thermodynamic observables, as well as the density of states and its fluctuations, in a clean Fermi liquid. Let us assign a scale dimension $[L] = -1$ to lengths, or $[k] = 1$ to wave numbers, and a scale dimension $[E] = [\omega] = z$ to energy and frequency; that is, the frequency scales with the wave number as $\omega \sim k^z$ [21]. $[L]$ is thus fixed by a convention, whereas the value of z depends on the nature of the soft modes in the system, and in general there may be more than one scale dimension z (see following). Now consider an observable A that depends on the wave number k, the frequency ω, the temperature T, and the magnetic field H, whose scale dimension $[H]$ depends on the physical situation. In general, A will consist of a regular or nonscaling part, and a scaling part δA with a scale dimension $[A]$ that is given by the naive or engineering dimension of A [22]. The scaling hypothesis states that δA obeys a homogeneity law

$$\delta A(k,\omega,T,H) = b^{-[A]}\,\delta A(kb,\omega b^z,Tb^z,Hb^{[H]}), \quad (2.1)$$

where $b > 0$ is the (arbitrary) length rescaling factor. The nonscaling part does not satisfy such a homogeneity law.

If there is more than one class of soft modes in a system, then there will be more than one dynamical exponent z. This is more common at critical fixed points than at stable ones, and for much of our discussion there will be only one z. However, for electrons interacting via a long-ranged Coulomb interaction, the plasmon is soft for all dimensions $d < 3$, and it scales differently than the other soft modes, so there are two separate dynamical exponents. As long as the scaling functions in question are regular functions of their arguments in the limit of small arguments, the largest of these various z will determine the leading nonanalytic behavior of the observable. However, there are important exceptions to this rule if the scaling function is *not* a regular function of all of its arguments. The DOS in the presence of plasmons represents an example, as we will see below.

We finally note that the choice of variables one assigns a nonzero scale dimension to is physically motivated, and depends on the fixed point under consideration. Generally, variables that take the system away from the fixed point under consideration, or take it from one fixed point to another one with a different symmetry, carry a positive scale dimension. An example is the magnetic field. By contrast, a change of the chemical potential μ takes the system from a Fermi-liquid fixed point to an equivalent Fermi-liquid fixed point with the same symmetry. The chemical potential is therefore assigned a scale dimension of zero, $[\mu] = 0$, although its naive dimension is that of an energy. This is a simple, but important point: The scaling hypothesis involves much more than just dimensional analysis.

A. Soft modes and universality classes

We are interested in universal behavior at low temperature in the limit of small wave numbers and small frequencies that

is caused by the presence of soft modes. The first question therefore needs to be about the nature of the soft modes in a Fermi liquid.

There are two types of modes that are soft at $T = 0$. The first class are single-particle excitations represented by the Green function $\langle\bar{\psi}_1\,\psi_2\rangle$, where $\bar{\psi}_1$ and ψ_2 are fermionic fields with labels $1 \equiv (\boldsymbol{x}_1,\tau_1,\sigma_1)$, $2 \equiv (\boldsymbol{x}_2,\tau_2,\sigma_2)$, etc., that comprise position $\boldsymbol{x}$, imaginary time τ, and spin projection σ. These excitations are soft at $T = 0$ because of the existence of a Fermi surface. They have a linear frequency-momentum relation (with the momentum measured from the Fermi wave number), and are effectively one dimensional since only excitations perpendicular to the $(d-1)$ dimensional Fermi surface are relevant. We will refer to them as the fermionic excitations. They determine the leading scaling behavior in a Fermi liquid, and they play a central role in Shankar's RG derivation of Landau Fermi-liquid theory [15].

Much more important for our purposes is a second class of soft modes. These are two-particle excitations of the type $\langle\bar{\psi}_1\,\bar{\psi}_2\,\psi_3\,\psi_4\rangle$. Since bilinear products of fermionic fields ψ and their adjoints $\bar{\psi}$ commute with each other as well as with fermion fields, these are effectively bosonic excitations, and we will refer to them as such. They in turn fall into two distinct classes. The first one are the familiar particle-hole-continuum excitations [23] and the corresponding excitations in the particle-particle or Cooper channel. They all have a linear frequency-momentum relation, and they appear in all angular momentum channels; in the s wave or $\ell = 0$ channel suitable linear combinations of them constitute the number density, spin density, and particle-particle density fluctuations. They are responsible, *inter alia*, for the familiar structure of the Lindhard function with its linear frequency-momentum scaling [23]. The second class comprises collective excitations. In the particle-hole spin-singlet channel, these include the excitations known as zero-sound modes in a neutral system, and the plasmons in a charged system. In the particle-hole spin-triplet channel, they are the paramagnon excitations. Both the zero-sound modes and the paramagnons also have a linear frequency-momentum relation, but their origin and physical nature is very different from the first class. They are the result of conservation laws (particle number and spin conservation, respectively) that guarantee their masslessness. The particle-hole continuum excitations, on the other hand, can be understood as the Goldstone modes of a spontaneously broken continuous symmetry that can be represented as a rotational symmetry between retarded and advanced degrees of freedom (see Ref. [19] and Sec. III A). Their softness is therefore not accidental either; it is controlled by a Ward identity and protected by Goldstone's theorem. As we will see, they provide the leading corrections to scaling in a Fermi liquid, and the effective field theory of Ref. [19] is formulated in terms of them. It is important to note that, due to the intrinsic coupling between the dynamics and the statics in a quantum system, these bosonic excitations lead, via mode-mode coupling effects, to long-ranged static correlations even in equilibrium systems. The same is not true for the fermionic excitations. These long-ranged correlations in turn fundamentally modify the nature of various quantum phase transitions [4,24,25].

The fact that all of the soft modes mentioned above have linear frequency-momentum relations is accidental. In systems with a long-ranged Coulomb interaction there is the plasmon excitation, whose frequency scales as the square root of the wave number in two-dimensional systems, and as a constant in three-dimensional ones. Another example of bosonic excitations whose frequency-momentum relation is different from that of the fermionic ones are the diffusive modes in a disordered electron system [7]. A technical consequence in all of these cases is the existence of more than one dynamical scale dimension z.

The Goldstone modes in the particle-hole spin-triplet channel and in the particle-particle channel are sensitive to an external magnetic field. The orbital effects of the field give a mass to the particle-particle channel, whereas the Zeeman effect gives a mass to two out of three modes in the particle-hole spin-triplet channel. We therefore have to distinguish between two universality classes: (1) A generic class with no symmetry-breaking fields, where all channels are soft, and (2) a magnetic-field universality class, where the Cooper channel is missing from the soft-mode spectrum, and only one mode in the particle-hole spin-triplet channel (the longitudinal one) is soft [26]. For the DOS and its susceptibility we further have to distinguish between the cases of a short-ranged interaction and a long-ranged Coulomb interaction, as we will see in Sec. II E.

B. Scaling of the free energy

Let us now consider the free energy density, from which all thermodynamic quantities can be derived (see below for the relation between the density of states and the free energy). In principle, theories formulated in terms of fermionic degrees of freedom only, such as Ref. [15], and theories formulated in terms of bosonic variables, such as Ref. [19], both contain the effects of both types of soft modes. However, in practice it is very difficult to extract the effects of one class of soft modes from a theory that has been formulated in terms of the other, and the effects of the soft modes in the limit of long wavelengths and low frequencies are additive. It therefore is natural to postulate the existence of two additive scaling contributions to the free energy density f, which we denote by $f^{(f)}$ and $f^{(b)}$, respectively, with the superscripts referring to the fermionic and bosonic soft modes, respectively [27]. Both depend on the temperature T and the magnetic field H. In addition, they depend on a source field h that is conjugate to the density of states. h is a generalized, frequency-dependent (for simplicity we do not show the frequency dependence explicitly) chemical potential, and the DOS can be interpreted as the order parameter of the broken symmetry (see Refs. [17,19] and Sec. III A). For the scaling part of the free energy density we thus write

$$f(T,H,h) = f^{(f)}(T,H,h) + f^{(b)}(T,H,h), \qquad (2.2)$$

and we postulate that $f^{(f)}$ and $f^{(b)}$ obey separate homogeneity laws.

We now need to determine the scale dimensions of $f^{(f)}$ and $f^{(b)}$, and their arguments. To this end we observe that, dimensionally, $f^{(f)}$ and $f^{(b)}$ are both an energy divided by a volume. However, for the fermionic excitations only the direction perpendicular to the Fermi surface is relevant [15], and the scale dimension of the fermionic part should therefore

be

$$[f^{(f)}] = 1 + z. \tag{2.3a}$$

For the bosonic part, we expect

$$[f^{(b)}] = d + z, \tag{2.3b}$$

with d the spatial dimensionality of the system. The temperature has a scale dimension $[T] = z$ (see above). For both the fermionic and bosonic modes the frequency scales as the wave number, and we therefore put

$$[T] = z = 1. \tag{2.3c}$$

We note, however, that in the case of a Coulomb interaction there is a second time scale set by the plasmon excitation, and hence a second dynamical exponent z [see the remarks in Sec. II A and after Eq. (2.1)]. We will explicitly deal with this in the context of Eqs. (2.11) and (2.19). We will not consider orbital effects of the magnetic field, i.e., we consider H only as it enters a Zeeman term, and hence

$$[H] = 1. \tag{2.3d}$$

Finally, h is a generalized chemical potential and therefore dimensionally an energy, so we also have

$$[h] = 1. \tag{2.3e}$$

However, the physical chemical potential carries a scale dimension of zero, see the remark at the end of the introduction to the current section,

$$[\mu] = 0. \tag{2.3f}$$

With these definitions, the DOS is given by $N = (\partial f/\partial h)/T$, which correctly makes N an inverse energy times an inverse volume. Combining all of these considerations, we now have

$$f^{(f)}(T,H,h,\mu) = b^{-(1+1)}\, f^{(f)}(Tb,Hb,hb,\mu), \tag{2.4a}$$

$$f^{(b)}(T,H,h,\mu) = b^{-(d+1)}\, f^{(b)}(Tb,Hb,hb,\mu). \tag{2.4b}$$

We note that there is an important physical difference between the physical fields H and T, and the derivatives of f with respect to them (i.e., thermodynamic quantities) on one hand, and the field h, which just serves as a source term that can not by physically realized (see Sec. III A) on the other. A related point is that the derivatives of f with respect to h, viz., the DOS and its susceptibility, are not gauge invariant quantities and hence show scaling behavior that is sensitive to a long-ranged Coulomb interaction, whereas the thermodynamic quantities show the same behavior for both short-ranged and Coulomb interactions. See Sec. IV B for a discussion of this point.

Before we analyze these homogeneity laws, let us illustrate an alternative way to determine the effective scale dimension of $f^{(f)}$. Suppose we do not use the above argument about effectively 1-d fermionic excitations, and assume instead that $f^{(f)}$ scales as an energy divided by a volume. However, we do acknowledge that the free energy depends on the microscopic wave number k_{F} and the microscopic energy ϵ_{F}, and use as input that $f^{(f)}$ is proportional to $k_{\mathrm{F}}^d/\epsilon_{\mathrm{F}}$ [28]. We thus write

$$f^{(f)}(T,h,H,\mu) = \frac{k_{\mathrm{F}}^d}{\epsilon_{\mathrm{F}}}\, \mathcal{F}(T,h,H,\mu). \tag{2.5a}$$

Then $\mathcal{F}$ is dimensionally an energy squared, and the scaling hypothesis is

$$\mathcal{F}(T,H,h,\mu) = b^{-2}\, \mathcal{F}(Tb,Hb,hb,\mu), \tag{2.5b}$$

which is equivalent to Eq. (2.4a). Knowledge about the dependence of $f^{(f)}$ on the microscopic parameters is thus equivalent to knowledge about the effective one dimensionality of the fermionic excitations. In Sec. IV A we will revisit these arguments and show how one can write a single free energy contribution that yields both the fermion and boson contributions by using the dependence of f on irrelevant operators.

C. Observables as derivatives of the free energy

The observables considered in this paper can be expressed in terms of derivatives of the free energy as follows. The DOS is given by the first derivative of f with respect to h,

$$N = \frac{1}{T}\left(\frac{\partial f}{\partial h}\right)_{h=0}, \tag{2.6a}$$

and the density-of-states susceptibility by the second one,

$$\chi_N = \frac{1}{T}\left(\frac{\partial^2 f}{\partial h^2}\right)_{h=0}. \tag{2.6b}$$

The entropy density s is the first derivative of f with respect to T,

$$s = \partial f/\partial T, \tag{2.7a}$$

and the specific-heat coefficient $\gamma_V = C_V/T$ is the second one,

$$\gamma_V = \partial^2 f/\partial T^2. \tag{2.7b}$$

The magnetization m, defined as the magnetic moment per unit volume, is the first derivative of f with respect to H,

$$m = \partial f/\partial H, \tag{2.8a}$$

and the spin susceptibility is the second one,

$$\chi_s = \partial^2 f/\partial H^2. \tag{2.8b}$$

Note that differentiating with respect to H yields the spin susceptibility, rather than the full magnetic susceptibility, since we have restricted the H dependence of the Hamiltonian to a Zeeman term. Also of interest is the (p-wave) nematic spin susceptibility that is given as the second derivative with respect to a field $\mathcal{H}$ that couples to the spin current rather than to the spin density,

$$\chi_{s,p\text{-wave}} = \partial^2 f/\partial \mathcal{H}^2. \tag{2.8c}$$

Finally, to make a point about naive dimensions and scaling functions, we also consider the density susceptibility,

$$\chi_{\mathrm{n}} = \partial n/\partial \mu = \partial^2 f/\partial \mu^2, \tag{2.9}$$

with n the particle number density. For reasons discussed below, χ_n scales differently than the other susceptibilities.

All of the above quantities have a naive dimension of inverse volume times inverse energy, except for χ_N, which is an inverse volume times an inverse energy squared. We will eventually conclude that in the case of short-ranged electron interaction, where there is only one dynamical exponent z, all of them except for χ_n have scaling properties that are simply given by their naive dimension. χ_n, however, does not scale according to its naive dimension. The physical reason is that changing the chemical potential does not change the symmetry of the system, nor does it move the system away from the stable, zero-temperature Fermi-liquid fixed point. That is, μ is not a RG-relevant operator with respect to the Fermi-liquid fixed point. This conclusion is confirmed by explicit perturbative calculations [29] .

D. Leading Fermi-liquid behavior

The formulation of Fermi-liquid theory in terms of a stable RG fixed point has been studied in great detail [15,30–33]. Here we give some simple arguments to show that our treatment is consistent with the existence of a stable Fermi-liquid fixed point, but we will not deal with the full complexity of the Fermi-liquid state. Our main objective will be the study of the leading corrections to scaling at this fixed point, which result from the bosonic soft modes. In this context, the Fermi-liquid results appear as a regular background for the nonanalyticities due to the bosonic soft modes and the resulting long-range correlations.

Let us consider the contributions of the fermionic excitations to the various observables. From Eqs. (2.6a) and (2.4a) or, alternatively, (2.5), we see that the DOS scales as a constant $N \sim$ const. Putting $b = 1/T$ in Eq. (2.4a) we see that the entropy density scales as $s \sim T$, and the specific-heat coefficient scales as $\gamma_V \sim$ const. For the magnetization, putting $b = 1/H$ in Eq. (2.4a) yields $m \sim H$, and hence the spin susceptibility also scales as a constant $\chi_s \sim$ const. The same is true for the nematic susceptibility $\chi_{s,p-\text{wave}}$ (and all higher nematic susceptibilities as well).

Using Eq. (2.9) and either Eq. (2.4a) or (2.5b), we see that $\chi_n \sim$ const as well. The scaling assumption thus reproduces the well-known properties of a Fermi liquid: The DOS and the susceptibilities γ_V, χ_s, and χ_n all are proportional to the bare DOS at the Fermi surface N_F. Notice, however, that the origin of χ_n scaling as a constant is very different from the analogous statements for the other susceptibilities, as it is a consequence of $[\mu] = 0$. The fact that all of these susceptibilities are trivially proportional to N_F by dimensional analysis thus masks an important difference. We will come back to this point below. The fields H and h do not couple to the fermionic excitations in any interesting way, and the latter therefore do not produce any nonanalytic corrections to the leading behavior.

Finally, for the DOS susceptibility χ_N, Eqs. (2.4a) and (2.6b) suggest $\chi_N \sim 1/T$. This is misleading, however. Repeated differentiations with respect to the conjugate field h at the same frequency do not lead to any frequency mixing, and as a consequence the Fermi-liquid result for χ_N is not as singular as naive scaling suggests; the second differentiation with respect to h in Eq. (2.6b) just produces a factor of $1/\epsilon_F$. In fact, the wave number (k) and frequency (ω) dependent generalization of χ_N at $T = 0$ for free electrons is $\chi_N^{(0)} \propto N_F/\epsilon_F$. Fermi-liquid corrections do not change the scaling of this result, and therefore we have $\chi_N \sim$ const. As we will see in the next subsection, for this quantity the leading contribution from the bosonic fluctuations in $d \leqslant 2$ is actually *stronger* than the Fermi-liquid contribution.

E. Leading corrections to Fermi-liquid scaling

Now consider the contributions of the bosonic excitations to the scaling part of the free energy, which yield the leading nonanalytic dependence of variables observables on the frequency, temperature, etc., and are not included in Landau Fermi-liquid theory. We will discuss the DOS, its susceptibility, the specific-heat coefficient, and the magnetic susceptibility, and distinguish between the cases of a short-range interaction and a long-range Coulomb interaction as appropriate. We note that the leading corrections to scaling can be large effects that are very important for an understanding of strongly correlated metals, and even have the potential for destroying the underlying Fermi-liquid state [17].

1. *Density of states*

The leading correction to the constant Fermi-liquid DOS is given by $\delta N = (\partial f^{(b)}/\partial h)/T|_{T=h=0}$. Let us first consider the case of a short-ranged interaction, in which the complications due to the presence of multiple time scales mentioned above are not relevant.

Short-ranged case. From Eq. (2.4b) we have

$$\delta N(\omega,T) = b^{1-d}\,\delta N(\omega b, Tb). \tag{2.10a}$$

This implies [34]

$$\delta N(\omega, T=0) \propto |\omega|^{d-1} \tag{2.10b}$$

and

$$\delta N(\omega=0,T) \propto T^{d-1}, \tag{2.10c}$$

and more generally

$$\delta N(\omega,T) = |\omega|^{d-1}\,F_N(T/|\omega|), \tag{2.10d}$$

where $F_N(x) = \delta N(1,x)$ is a scaling function. That is, the DOS at $T = 0$ is a nonanalytic function of the energy distance ω from the Fermi surface, and the DOS at the Fermi surface is a nonanalytic function of the temperature. For $1 < d < 3$ the leading nonanalyticity is stronger than the leading analytic correction, which is ω^2 or T^2, respectively. Note that scaling by itself does not guarantee that the prefactor of the nonanalyticity is nonzero in any given system in any given dimension. Also, the presence of dangerous irrelevant variables may invalidate the simple scaling assumption represented by Eq. (2.10a) [2]. However, barring such exceptional circumstances, the exponent in Eq. (2.10b) is expected to be exact. This is in sharp contrast to perturbative considerations, which can never guarantee that a stronger nonanalyticity will not appear at some higher order of perturbation theory. The remaining question is the validity of the scaling assumption. Establishing this will be the purpose of Sec. III, where we will derive the scaling behavior from renormalization-group arguments.

This will establish that the exponent in Eq. (2.10b) is indeed exact. The determination of the prefactor requires an explicit calculation. In Ref. [35] we will present a one-loop calculation that is not perturbative in the interaction strength and shows that the prefactor is generically nonzero. However, the case $d = 2$ is an exception; to one-loop order the prefactor of the $|\omega|$ vanishes, in agreement with previous results from many-body perturbation theory [36,37]. The leading nonanalytic contribution at one-loop order for $d = 2$ is an $|\omega|/\ln^3|\omega|$ that originates from the particle-particle channel [38]. The explicit loop expansion also shows that in special dimensions there are logarithmic corrections to power-law scaling, which the scaling theory is not sensitive to. For instance, in $d = 3$ we find $\delta N \propto \omega^2 \ln|\omega|$.

Long-range case.. As we will see in Secs. III and IV, the simple scaling arguments given above do indeed yield the correct results in the case of a short-ranged interaction between the electrons. However, in the case of a long-ranged Coulomb interaction in $d \leqslant 2$ the leading singularity of the DOS is even stronger. The reason is the presence of the plasmon time scale, whose frequency scales as $\omega \sim k^{(3-d)/2}$. In addition to the frequency scale with $z = 1$, which reflects the particle-hole excitations, we thus have a second frequency scale with $z = z_p = (3-d)/2$. As a result, Eq. (2.10a) gets generalized to

$$\begin{aligned}\delta N(\omega) &= b^{1-d}\,\delta N(\omega b, \omega b^{(3-d)/2}) \\ &= |\omega|^{d-1}\, f_N(\omega^{(d-1)/2}),\end{aligned} \tag{2.11}$$

where $f_N(x) = \delta N(1,x)$ is a scaling function and we have dropped the temperature dependence for simplicity. Naively, one would expect $f_N(x \to 0) = \text{const}$, which would lead to Eq. (2.10b). However, it turns out that the subleading frequency scale characterized by z_p is a dangerous irrelevant variable for the DOS if $d < 2$, namely, $f_N(x \to 0) \propto x^{-2(2-d)/(3-d)} + O(1)$. As a result,

$$\delta N(\omega) \propto \begin{cases} |\omega|^{(d-1)/(3-d)} & \text{for } 1 < d < 2, \\ |\omega|^{d-1} & \text{for } 2 \leqslant d < 3. \end{cases} \tag{2.12}$$

Note that the existence of a dangerous irrelevant variable can not be deduced from scaling arguments alone; establishing the behavior described above requires an explicit calculation (see Refs. [36,37] and Appendix D). In $d = 3$ the leading behavior is again $\omega^2 \ln|\omega|$, as in the short-range case.

2. *Density-of-states susceptibility*

We next recall what scaling arguments predict for the temperature and frequency dependence of the density-of-states fluctuations [39]. Again, we first consider the short-range case.

Short-range case. Let us generalize Eq. (2.10a) by keeping the auxiliary source field h conjugate to the DOS. We then obtain

$$\delta N(\omega, T, h) = b^{(1-d)}\,\delta N(\omega b, Tb, hb). \tag{2.13}$$

Differentiating again with respect to h, and setting $h = 0$, we obtain a homogeneity law for the DOS susceptibility

$$\chi_{\rm N}(k,\omega,T) = b^{2-d}\,\chi_{\rm N}(kb,\omega b,Tb), \tag{2.14}$$

where we have added a dependence on the wave number k. Alternatively, we can consider the homogeneous susceptibility in a system with a finite linear dimension L. All homogeneity laws then remain valid with k replaced by $1/L$. Since frequency and temperature scale the same way, we set $\omega = 0$ for simplicity. For the static susceptibility as a function of k and T Eq. (2.14) implies

$$\chi_{\rm N}(k,T) = k^{d-2}\, f_\chi(T/v_{\rm F}k), \tag{2.15}$$

with f_χ a scaling function and $v_{\rm F}$ the Fermi velocity. At this point we need to acknowledge that $\chi_{\rm N}$, because of the frequency structure of the underlying four-fermion correlation function, has two intrinsically different parts, one of which is constant as $T \to 0$, whereas the other vanishes linearly with T. This is explained in Appendix E, and it is consistent with general statistical arguments (see Ref. [39] and Sec. IV E). f_χ, and hence $\chi_{\rm N}$, therefore come with two scaling parts, viz.,

$$\chi_{\rm N}^{(0)}(k,T) = \frac{1}{k^{2-d}}\, f_\chi^{(0)}(T/v_{\rm F}k), \tag{2.16a}$$

$$\chi_{\rm N}^{(1)}(k,T) = \frac{T}{k^{3-d}}\, f_\chi^{(1)}(T/v_{\rm F}k), \tag{2.16b}$$

with $f_\chi^{(0)}(x \to 0) = \text{const}$ and $f_\chi^{(1)}(x \to 0) = \text{const}$. In the limit $T \ll v_{\rm F}k$, $\chi_{\rm N}^{(1)}$ is thus small compared to $\chi_{\rm N}^{(0)}$ by a factor of $T/v_{\rm F}k$. In the opposite limit $v_{\rm F}k \ll T$, both parts yield

$$\chi_{\rm N}(k=0,T) \propto 1/T^{2-d}. \tag{2.17}$$

Of particular interest are the physical dimensions $d = 2$ and $d = 3$, where the zero exponents in Eqs. (2.16) and (2.17) signify logarithms. An explicit calculation [39] yields for the leading behavior in $d = 2$

$$\chi_{\rm N}(k,T) \propto \begin{cases} \ln(1/k) & \text{for } T \ll v_{\rm F}k, \\ \ln(1/T) & \text{for } v_{\rm F}k \ll T \end{cases} \tag{2.18a}$$

from $\chi_{\rm N}^{(0)}$, and in $d = 3$,

$$\chi_{\rm N}(k,T) \propto \begin{cases} k\,[1 + (T/v_{\rm F}k)\ln(1/k)] & \text{for } T \ll v_{\rm F}k, \\ T\ln(1/T) & \text{for } v_{\rm F}k \ll T \end{cases} \tag{2.18b}$$

with the logarithms coming from $\chi_{\rm N}^{(1)}$.

This highly nonanalytic behavior of the DOS susceptibility reflects the same correlations that lead to the nonanalyticity in the DOS itself [Eq. (2.10b)]. Again, the exponents are expected to be exact.

Long-range case. Similar to the case of the DOS, the DOS susceptibility gets modified by a Coulomb interaction, but for this observable the plasmon frequency scale is a dangerous irrelevant variable even for $d > 2$. Equation (2.15) gets generalized to

$$\begin{aligned}\chi_{\rm N}(k,\omega=0,T) &= k^{d-2} f_\chi\left(\frac{T}{v_{\rm F}k}, \frac{T}{v_{\rm F}\kappa^{(d-1)/2}\,k^{(3-d)/2}}\right) \\ &= k^{d-2} g_\chi[T/k, (k/\kappa)^{(d-1)/2}],\end{aligned} \tag{2.19}$$

where κ is the screening wave number and g_χ is another scaling function. $g_\chi(x,y)$ again has two separate scaling parts, for the reasons explained above. In the long-range case, the one proportional to T always gives the leading contribution, which

generalizes Eq. (2.16b):

$$\chi_{\rm N}(k,T) = \frac{T}{k^{3-d}}\, g_\chi^{(1)}[T/k,(k/\kappa)^{(d-1)/2}]. \tag{2.20}$$

If $g_\chi^{(1)}(x,y)$ were a regular function of its second argument for $y \to 0$, then the leading behavior of χ_N would again be given by Eq. (2.16b) since the plasmon frequency scale is subleading compared to the ballistic scale, and thus irrelevant in the renormalization-group sense. However, this subleading scale is again dangerously irrelevant. In order to see this one needs to perform an explicit calculation (see Ref. [39] and Appendix D); here we just list the results. The dimensionality dependence is complicated, and we focus on the physical dimensions $d = 2,3$. In $d = 2$ one finds

$$\chi_{\rm N}(k,T) \propto \kappa^2 T \times \begin{cases} (1/T^3)\,\ln\left(T^2/v_{\rm F}^2\kappa k\right) & \text{for } v_{\rm F}k \ll T^2/v_{\rm F}\kappa, \\ 1/(v_{\rm F}\kappa)^{3/2}T^{3/2} & \text{for } T^2/v_{\rm F}\kappa \ll v_{\rm F}k \ll T, \\ 1/(v_{\rm F}\kappa)^{3/2}(v_{\rm F}k)^{3/2} & \text{for } T \ll v_{\rm F}k. \end{cases} \tag{2.21}$$

In $d = 3$ the result is

$$\chi_{\rm N}(k,T) \propto \kappa^3 T/(v_{\rm F}\kappa)^2 v_{\rm F}k \tag{2.22}$$

for all values of T and k. Note that χ_N diverges for $k \to 0$ at fixed T in both $d = 2$ and 3. We will discuss the significance of this result in Sec. IV.

3. *Specific-heat coefficient*

We next consider the specific-heat coefficient. The homogeneity law for the leading correction to the constant Fermi-liquid contribution $\delta\gamma_V = \partial^2 f^{(b)}/\partial T^2$ is

$$\delta\gamma_V(T,H) = b^{-(d-1)}\,\delta\gamma(Tb,Hb). \tag{2.23}$$

In a zero field, this results in

$$\delta\gamma_V(T) \propto T^{d-1}. \tag{2.24a}$$

In a nonzero magnetic field, this gets generalized to

$$\delta\gamma_V = T^{d-1}\, g_\gamma(H/T), \tag{2.24b}$$

with g_γ a scaling function. These results are all consistent with perturbation theory [37,40].

4. *Spin susceptibility*

Now we consider the leading correction to the spin susceptibility $\delta\chi_{\rm s} = \partial^2 f^{(b)}/\partial H^2$. This is due to soft modes in the spin-triplet channel, and thus is the same for the short- and long-range cases. From Eq. (2.4b), and including a wave-number dependence of $\chi_{\rm s}$, we find

$$\delta\chi_{\rm s}(k,T,H) = b^{-(d-1)}\,\delta\chi_{\rm s}(kb,Tb,Hb). \tag{2.25}$$

For the wave-number-dependent spin susceptibility at $T = 0$ and $H = 0$ this yields

$$\delta\chi_{\rm s} \propto k^{d-1}. \tag{2.26}$$

This is consistent with perturbative results [29,41,42], but the scaling arguments are not sensitive to a logarithmic term in $d = 3$, where the behavior is $k^2 \ln k$ [29]. The nonanalytic T and H dependencies of the homogeneous spin susceptibility are

$$\delta\chi_{\rm s} \propto T^{d-1} \tag{2.27a}$$

and

$$\delta\chi_{\rm s} \propto H^{d-1}, \tag{2.27b}$$

respectively. These scaling results are also consistent with perturbation theory [29,40,43].

The scaling behavior of $\chi_{{\rm s},p-{\rm wave}}$is the same as that of $\chi_{\rm s}$, except that H gets replaced by $\mathcal{H}$. This is important in the context of the quantum phase transition from a paramagnet to a spin-nematic phase, and renders the transition generically first order [24]. We also note that $\chi_{{\rm s},p-{\rm wave}}$ is a nonanalytic function of $\mathcal{H}$, but not of H, since $\mathcal{H}$ does not couple to the spin density. Analogous statements hold for other susceptibilities and fields: Scaling arguments give only the functional dependence of observables on fields; a nonzero prefactor requires, among other things, a nonvanishing coupling.

F. Signs of the leading corrections

The scaling arguments presented in this section give information about the various power laws that characterize the leading nonanalyticities, but make no statement about the prefactors. However, with some additional physical reasoning one can give strong arguments at least for what the signs of the various effects should be, which then can be confirmed by explicit calculations.

Let us start with the spin susceptibility. Its nonanalyticity is a result of fluctuations about Stoner theory that weaken the tendency towards ferromagnetism. As a result, $\chi_{\rm s}(k = 0,T = 0)$ will decrease. A nonzero k or T in turn weakens the soft-mode effect, therefore the sign of the leading nonanalyticity in Eqs. (2.26) and (2.27a) will be positive. This is indeed the result first obtained in perturbation theory in Ref. [29], which has profound consequences for the ferromagnetic quantum phase transition [4,25,44]. We note that in disordered systems the signs of the corresponding nonanalyticities are opposite [45]. The reason is that disorder slows down the electrons (diffusive rather than ballistic motion), which effectively enhances the interaction and hence the tendency towards magnetism. This effect in turn is weakened by a nonzero k or T. The prefactor in Eq. (2.27b) is also positive since a magnetic field enhances the tendency toward magnetism. This is also in agreement with prior perturbative results [40].

The signs of the corrections to the DOS and the specific-heat coefficient can be understood as follows. The correlations induced by the Goldstone modes lead to long-range correlations that tend to order the system. The entropy is thus expected to decrease as a result of them, and so will the specific heat. We thus expect the prefactor in Eqs. (2.24) to be negative, which is indeed borne out by explicit calculations [37,40]. This is consistent with the sign of the DOS correction (2.10b), which is also known to be negative [36,46].

III. RG-BASED DERIVATION OF SCALING

In this section we show how the above results can be derived without invoking a scaling assumption, by performing a RG

analysis of the effective field theory of Ref. [19]. We stress again that even though the Fermi-liquid fixed point is not a critical fixed point, it nevertheless displays scale invariance due to the existence of Goldstone modes. Therefore, useful results for the entire Fermi-liquid phase can be obtained from very simple RG arguments. Furthermore, in a properly formulated theory of a stable phase, non-Gaussian terms are RG irrelevant and, as a consequence of this, *exact* scaling exponents can be simply obtained [47]. This is in contrast to the situation at a critical fixed point, where the explicit calculation of exponents usually involves an expansion in an artificial small parameter, such as the deviation from a critical dimension [1].

For the convenience of the reader, we first summarize the key arguments from Ref. [19] that lead to the identification of the soft modes (Sec. III A), and the structure of the field theory that separates the soft modes from the massive ones (Sec. III B). For details we refer the reader to that reference. In the remainder of this section we perform a RG analysis of the resulting effective field theory.

A. Broken symmetry and Goldstone modes in a Fermi liquid

The theory of Ref. [19] is formulated in terms of bosonic matrix fields $Q_{nm}(\boldsymbol{x},\boldsymbol{y})$ that are isomorphic to bilinear products of fermion fields according to

$$Q_{nm}(\boldsymbol{x},\boldsymbol{y}) \cong \frac{i}{2}\begin{pmatrix} -\psi_{n\uparrow}(\boldsymbol{x})\bar{\psi}_{m\uparrow}(\boldsymbol{y}) & -\psi_{n\uparrow}(\boldsymbol{x})\bar{\psi}_{m\downarrow}(\boldsymbol{y}) & -\psi_{n\uparrow}(\boldsymbol{x})\psi_{m\downarrow}(\boldsymbol{y}) & \psi_{n\uparrow}(\boldsymbol{x})\psi_{m\uparrow}(\boldsymbol{y}) \\ -\psi_{n\downarrow}(\boldsymbol{x})\bar{\psi}_{m\uparrow}(\boldsymbol{y}) & -\psi_{n\downarrow}(\boldsymbol{x})\bar{\psi}_{m\downarrow}(\boldsymbol{y}) & -\psi_{n\downarrow}(\boldsymbol{x})\psi_{m\downarrow}(\boldsymbol{y}) & \psi_{n\downarrow}(\boldsymbol{x})\psi_{m\uparrow}(\boldsymbol{y}) \\ \bar{\psi}_{n\downarrow}(\boldsymbol{x})\bar{\psi}_{m\uparrow}(\boldsymbol{y}) & \bar{\psi}_{n\downarrow}(\boldsymbol{x})\bar{\psi}_{m\downarrow}(\boldsymbol{y}) & \bar{\psi}_{n\downarrow}(\boldsymbol{x})\psi_{m\downarrow}(\boldsymbol{y}) & -\bar{\psi}_{n\downarrow}(\boldsymbol{x})\psi_{m\uparrow}(\boldsymbol{y}) \\ -\bar{\psi}_{n\uparrow}(\boldsymbol{x})\bar{\psi}_{m\uparrow}(\boldsymbol{y}) & -\bar{\psi}_{n\uparrow}(\boldsymbol{x})\bar{\psi}_{m\downarrow}(\boldsymbol{y}) & -\bar{\psi}_{n\uparrow}(\boldsymbol{x})\psi_{m\downarrow}(\boldsymbol{y}) & \bar{\psi}_{n\uparrow}(\boldsymbol{x})\psi_{m\uparrow}(\boldsymbol{y}) \end{pmatrix}. \tag{3.1}$$

Here $\psi_{n\sigma}(\boldsymbol{x})$ and $\bar{\psi}_{n\sigma}(\boldsymbol{x})$ are fermionic fields that depend on a real-space position $\boldsymbol{x}$, a frequency index n that represents a fermionic Matsubara frequency $\omega_n = 2\pi T(n+1/2)$, and a spin projection $\sigma = \uparrow,\downarrow$. The Green function $G_{n\sigma}(\boldsymbol{x}-\boldsymbol{y}) = \langle\psi_{n\sigma}(\boldsymbol{x})\,\bar{\psi}_{n\sigma}(\boldsymbol{y})\rangle$ can obviously be constructed from the matrix elements of the expectation value $\langle Q_{nm}(\boldsymbol{x},\boldsymbol{y})\rangle$. In particular, the density of states $N(\omega)$ is given as the spectrum of a causal function $Q(i\omega_n)$,

$$N(\omega) = \frac{1}{\pi}\mathrm{Re}\,\langle \mathrm{tr}\, Q_{nn}(\boldsymbol{x},\boldsymbol{x})\rangle|_{i\omega_n\to\omega+i0}, \tag{3.2}$$

where tr traces over the 4×4 matrix in Eq. (3.1). Any fermionic action can be rewritten in terms of the Q by implementing the isomorphism [Eq. (3.1)] by means of a Lagrange multiplier field, expressing all four-fermion terms in terms of the Q, and integrating out the fermions.

However, the resulting action in terms of the Q is not more useful than the original fermionic action unless one has a way of separating soft and massive modes. To this end we consider a continuous transformation in frequency space that takes the form, for a fixed pair of frequencies n_1, n_2,

$$Q_{n_1n_2}(\boldsymbol{x},\boldsymbol{y}) \to Q_{n_1n_2}(\boldsymbol{x},\boldsymbol{y}) + \delta Q_{n_1n_2}(\boldsymbol{x},\boldsymbol{y}), \tag{3.3a}$$

where

$$\delta Q_{n_1n_2}(\boldsymbol{x},\boldsymbol{y}) = \int dz\,[\varphi(\boldsymbol{x},z)\,Q_{n_2n_2}(z,\boldsymbol{y}) - Q_{n_1n_1}(\boldsymbol{x},z)\,\varphi(z,\boldsymbol{y})]. \tag{3.3b}$$

For a pair of the underlying fermionic fields with frequencies n_1 and n_2, respectively, this amounts to a transformation into linear combinations of the same pair, with a nonlocal mixing angle $\varphi(\boldsymbol{x},\boldsymbol{y})$. By considering how the action transforms under this transformation, one can use standard techniques [48] to derive a Ward identity that relates expectation values of Q, or two-fermion correlations, to higher correlation functions. For free fermions, only four-fermion correlations appear and the Ward identity takes the form

$$\begin{aligned}&\langle \mathrm{tr}\, Q_{n_1n_2}(\boldsymbol{p};\boldsymbol{k})\,\mathrm{tr}\, Q_{n_3n_4}(\boldsymbol{p}';\boldsymbol{q})\rangle \\ &\quad= \frac{i}{2}\,\delta_{\boldsymbol{p},\boldsymbol{p}'}\,\delta_{n_1n_3}\,\delta_{n_2n_4} \\ &\qquad\times \frac{\langle \mathrm{tr}\, Q_{n_1n_1}(\boldsymbol{p}-\boldsymbol{k}/2)\rangle - \langle \mathrm{tr}\, Q_{n_2n_2}(\boldsymbol{p}+\boldsymbol{k}/2)\rangle}{i\Omega_{n_1-n_2} + \boldsymbol{p}\cdot\boldsymbol{k}/m_\mathrm{e}}.\end{aligned} \tag{3.4}$$

Here $Q_{nm}(\boldsymbol{p};\boldsymbol{k}) = Q_{nm}(\boldsymbol{p}+\boldsymbol{k}/2,\boldsymbol{p}-\boldsymbol{k}/2)$ is a spatial Fourier transform of $Q_{nm}(\boldsymbol{x},\boldsymbol{y})$, m_e is the electron mass, and $\Omega_{n-m} = 2\pi T(n-m)$ is a bosonic Matsubara frequency. For vanishing wave vector $\boldsymbol{q}$ and frequency $\Omega_{n_1-n_2}$, the four-fermion correlation function represented by the left-hand side of Eq. (3.4) diverges, provided the difference of Green functions in the numerator on the right-hand side is nonzero in this limit. This is the case in the energy or frequency region that represents the support of the spectrum of the Green function, which coincides with the region where the density of states is nonzero, provided the two fermionic frequencies ω_{n_1} and ω_{n_2} have opposite signs. The degrees of freedom represented by $Q_{n_1n_2}$ with $\omega_{n_1}\omega_{n_2} < 0$ thus are soft modes. Physically, they are the Goldstone modes due to a spontaneous breaking of the continuous symmetry represented by Eqs. (3.3), i.e., a rotation in frequency spaces that mixes retarded and advanced degrees of freedom. The corresponding order parameter is the spectrum of the Green function, which is related to the DOS via Eq. (3.2).

An important concept in this context is the field conjugate to the order parameter. The correlation function in Eq. (3.2) can be generated by adding to the action a source term

$$\sum_n h_n \int d\boldsymbol{x}\, Q_{nn}(\boldsymbol{x},\boldsymbol{x}), \tag{3.5}$$

with h_n a frequency-dependent source field that relates to the order parameter of the Fermi liquid the same way an external magnetic field relates to the magnetization in a ferromagnet. It is the field we have denoted by h in Sec. II, dropping the frequency index. From Eqs. (3.5) and (3.1) we see that physically h_n is a generalized, frequency-dependent chemical potential. A frequency independent h_n would represent just

a shift of the chemical potential. The frequency dependence is crucial, however, as is obvious from the above discussion, since the zero-frequency limit depends on whether the real axis is approached from above or below.

The remaining question is how the electron-electron interaction affects these conclusions. An analysis of the Ward identity shows that additional terms appear in the numerator of the right-hand side of Eq. (3.4). Since these have a different dependence on the electron-electron interaction than the two-fermion terms in Eq. (3.4), they can not cancel, and the conclusion regarding the soft modes remains unchanged. This of course reflects one of the main tenets of Fermi-liquid theory, namely, that interactions do not qualitatively change the basic features of the noninteracting system.

To summarize, the fluctuations of the bilinear fermion field Q_{nm} shown in Eq. (3.1) are soft if $\omega_n\omega_m < 0$, and massive otherwise. These soft and massive modes were denoted by q_{nm} and P_{nm}, respectively, in Ref. [19], where it was also shown that a nonzero temperature gives the q_{nm} a mass.

B. Structure of the field theory

The theory of Ref. [19] is formulated in terms of the soft matrix field $q_{nm}(\boldsymbol{k}) \equiv \sum_{\boldsymbol{p}} q_{nm}(\boldsymbol{p};\boldsymbol{k})$ and the massive one $P_{nm}(\boldsymbol{k}) \equiv \sum_{\boldsymbol{p}} P_{nm}(\boldsymbol{p};\boldsymbol{k})$ [49] that were identified in Sec. III A and encode the soft and massive components of bilinear fermion fields $\bar{\psi}_n\psi_m$. The softness of the q is guaranteed by a Ward identity, Eq. (3.4). The order-parameter field Q_{nn} consists of a saddle-point contribution that contains the free-electron contribution and Fermi-liquid corrections to it, and a contribution that represents the bosonic fluctuations and is given by P_{nn}. The scaling part of the DOS that was denoted by δN in Sec. II E 1 then takes the form

$$\delta N(\omega) = \mathrm{Re}\, Q(i\omega_n \to \omega + i0) \tag{3.6a}$$

with

$$Q(i\omega_n) = \frac{1}{\pi}\frac{1}{V}\sum_{\boldsymbol{p}} \langle \mathrm{tr}\, P_{nn}(\boldsymbol{p})\rangle. \tag{3.6b}$$

The Goldstone modes affect δN since the massive modes P couple to the soft modes q (see below).

The effective action $\mathcal{A}$ takes the form of an expansion in powers of q and P [see Eqs. (4.45)–(4.47) in Ref. [19]] [50]. In a symbolic notation that shows only quantities that carry a scale dimension, viz., the fields $q_{nm}(\boldsymbol{k}) \equiv q$ and $P_{nm}(\boldsymbol{k}) \equiv P$, and factors of volume V, wave number k, and frequency ω (which we do not need to distinguish from factors of temperature for our purposes), the Gaussian action takes the form [see Eq. (4.45) in Ref. [19]]

$$\mathcal{A}^{(2)} = \frac{1}{V}\sum_{k,\omega}[k + \omega + \gamma\,\omega]\, q^2 + \frac{1}{V}\sum_{k,\omega}[1 + \gamma\omega]\, P^2. \tag{3.7}$$

Here and in what follows the sums are over the appropriate sets of wave vectors and frequencies, and the powers of k and ω in each term follow from the properties of the convolutions of Green's functions that make up the vertices of the theory in the limit of long wavelengths and small frequencies (see Ref. [19]). As mentioned above, ω can stand for either frequency or temperature. The factors k and ω in the vertices should be understood as being multiplied by functions of k/ω, which are of $O(1)$ for scaling purposes and are not shown for simplicity. γ represents the interaction amplitude. Notice that the interacting and noninteracting parts of the Gaussian q vertex both are linear in k or ω, whereas the interacting part of the P vertex carries a factor of ω compared to the noninteracting one.

The non-Gaussian terms $\Delta\mathcal{A}$ are given in terms of fields $\not{q}$ and $\not{P}$ that are closely related to q and P. Their operational definition is that their propagators are the q and P propagators, respectively, with the noninteracting parts subtracted [50]. From Eq. (3.7) we see that the q and $\not{q}$ propagators scale the same way, viz.,

$$\langle q\, q\rangle \sim \langle \not{q}\, \not{q}\rangle \sim \frac{V}{k+\omega}, \tag{3.8a}$$

whereas the P and $\not{P}$ propagators scale differently,

$$\langle P\, P\rangle \sim V \times \mathrm{const}, \tag{3.8b}$$

$$\langle \not{P}\, \not{P}\rangle \sim V\omega. \tag{3.8c}$$

The mixed propagators $\langle q\, \not{q}\rangle$ and $\langle P\, \not{P}\rangle$ are equal to $\langle q\, q\rangle$ and $\langle P\, P\rangle$, respectively. For scaling purposes we therefore need to distinguish between P and $\not{P}$, but not between q and $\not{q}$. With this in mind, $\Delta\mathcal{A}$ takes the form [see Eqs. (4.47) and (4.48) in Ref. [19]]

$$\Delta\mathcal{A} = \Delta\mathcal{A}^{(3)} + \Delta\mathcal{A}^{(4)} + \ldots, \tag{3.9}$$

where, in the same schematic notation as in Eq. (3.7),

$$\begin{aligned}\Delta\mathcal{A}^{(3)} = {} & \frac{c_{3,0}}{V^2}\sum_{\{k,\omega\}}[\gamma\,\omega + O(\gamma^3\omega^3)]q^3 \\ & + \frac{c_{2,1}}{V^2}\sum_{\{k,\omega\}}[1 + O(\gamma^2\omega^2)]q^2\not{P} \\ & + \frac{c_{1,2}}{V^2}\sum_{\{k,\omega\}}\gamma\,\omega\, q\not{P}^2 + \frac{c_{0,3}}{V^2}\sum_{\{k,\omega\}}\not{P}^3,\end{aligned} \tag{3.10a}$$

$$\begin{aligned}\Delta\mathcal{A}^{(4)} = {} & \frac{c_{4,0}}{V^3}\sum_{\{k,\omega\}}[k + \omega + \gamma^2\omega^2/k + O(\gamma^4\omega^4)]q^4 \\ & + \frac{c_{3,1}}{V^3}\sum_{\{k,\omega\}}[\gamma\omega/k + O(\gamma^3\omega^3)]q^3\not{P} \\ & + \frac{c_{2,2}}{V^3}\sum_{\{k,\omega\}}[1/k + O(\gamma^2\omega^2)]q^2\not{P}^2 \\ & + \frac{c_{1,3}}{V^3}\sum_{\{k,\omega\}}\gamma\omega\, q\not{P}^3 + \frac{c_{0,4}}{V^3}\sum_{\{k,\omega\}}\not{P}^4,\end{aligned} \tag{3.10b}$$

where the $c_{n,m}$ are coupling constants. Notice that various vertices in $\Delta\mathcal{A}$, e.g., the leading term of $O(q^2\not{P}^2)$, are singular functions of k (or ω) for small k. These singularities get stronger with increasing order in the fields; for instance, the leading term of $O(q^2\not{P}^{2n})$ has a vertex that scales as $1/k^{2n-1}$. As we will show below, these singular vertices do not pose a problem for our purposes. We stress again that this form of the action is highly schematic and can be used for power-counting purposes only; many features that are crucial of explicit calculations have been suppressed for clarity. See

Ref. [19] for a complete expression. We also note that the schematic notation ignores a structural difference between the particle-hole and particle-particle channels that is only logarithmic in nature and hence does not appear at the level of power counting. However, it is of qualitative importance once logarithmically small effects are taken into account (see the next subsection).

C. Leading scaling behavior, and the fixed-point action

Equation (3.7) accurately represents the schematic form of the Gaussian action in the particle-hole channel, which was the only one considered in Ref. [19]. In particular, it accurately represents the Gaussian propagator in the particle-hole channel, which has the schematic structure

$$\langle q\, q\rangle_{\text{p-h}} = \frac{V}{k + (1+\gamma)\omega}. \tag{3.11}$$

We will proceed by first analyzing the action in the particle-hole channel from a RG point of view, and then consider the particle-particle or Cooper channel. We also recall that the particle-particle channel is sensitive to a small magnetic field, which gives its soft modes a mass, and therefore can always be suppressed (see Sec. II A). This effect is qualitatively the same as in disordered electron systems, where the orbital effects of a small magnetic field suppress the diffusive modes known as Cooperons [51].

We now look for a fixed point of the action [Eqs. (3.7) and (3.10)] that describes a Fermi liquid. We use Ma's method of choosing scale dimensions for all relevant quantities and then showing self-consistently that these choices lead to a stable fixed point [2]. As in Sec. II we assign a scale dimension $[k] = 1$ to wave numbers, and $[\omega] = 1$ to frequencies (i.e., we choose a dynamical exponent $z = 1$). The latter choice reflects the linear dispersion relation of the soft modes [see the first term in Eq. (3.7)], which in a Fermi liquid we do not expect to be changed by renormalization. We further do not expect the power of wave number (or frequency) in the Gaussian vertex to be renormalized, and therefore assign a scale dimension $[q(\boldsymbol{k})] = -(d+1)/2$ and $[q(\boldsymbol{x})] = (d-1)/2$ to the soft field as a function of $\boldsymbol{k}$ and $\boldsymbol{x}$, respectively (i.e., we choose the exponent η to be zero). The P propagators, normalized by the volume, are expected to scale as constants, as they do at Gaussian order [see Eq. (3.8b)]. We hence assign a scale dimension $[P(\boldsymbol{k})] = -d/2$. $\not{P}$ scales differently [Eq. (3.8c)]. This behavior, which again is not expected to change under renormalization, implies $[\not{P}(\boldsymbol{k})] = -(d-1)/2$. It is important to stress that all of these expectations will be verified self-consistently once the RG scheme is complete, and do *not* constitute *ad hoc* assumptions.

With these choices, the q^2 term in Eq. (3.7) is dimensionless; in particular, $[\gamma] = 0$. The constant contribution to the P^2 term is also dimensionless, whereas the $\gamma\omega$ contribution is irrelevant by one power of wave number or frequency compared to the constant one. We now determine the scale dimensions of the non-Gaussian terms. For the coupling constants of the cubic terms we find $[c_{3,0}] = [c_{2,1}] = -(d-1)/2$, $[c_{1,2}] = [c_{0,3}] = -(d+3)/2$, and for those of the quartic ones $[c_{4,0}] = [c_{3,1}] = [c_{2,2}] = -(d-1)$, $[c_{1,3}] = [c_{0,4}] = -(d+2)$. All of the non-Gaussian terms thus have negative scale dimensions for all $d > 1$. It is easy to verify that this is also true for all terms of higher order in the fields, despite the singular vertices in $\mathcal{A}_{q-P}$ mentioned above. For instance, the $q^2\not{P}^4$ term, whose vertex scales as $1/k^3$, has a coupling constant with scale dimension $[c_{2,4}] = -2(d-1)$. At tree level, the fixed-point action is thus given by the Gaussian terms alone, and all others are irrelevant with respect to the Fermi-liquid fixed point in all dimensions $d > 1$. It follows by standard arguments [1] that this remains true order by order in a loop expansion. All coefficients will in general acquire finite renormalizations, but the structure of the theory will not change. An important ingredient in this chain of arguments is the Ward identity proven in Ref. [19], which identifies q as a soft mode. This assures that the q vertices will remain soft under renormalization.

We now consider the particle-particle or Cooper channel. The action is again schematically given by Eqs. (3.7), (3.9), and (3.10), but with one crucial difference: The frequency structure (which we have suppressed in our schematic notation) is different and leads, upon inversion of the quadratic form, to the characteristic Cooper-ladder structure of the Gaussian q propagator. Schematically one obtains, instead of Eq. (3.11),

$$\langle q\, q\rangle_{\text{p-p}} = \frac{V}{k+\omega} + \frac{V\gamma_{\text{c}}\,\omega}{(k+\omega)^2}\,\frac{1}{1+\gamma_{\text{c}}\,\ln(1/\omega)}\,, \tag{3.12}$$

where γ_{c} is the interaction amplitude in the Cooper channel. As before, we do not distinguish between factors of frequency and factors of temperature. The structure of the interaction part of Eq. (3.12) implies that the q fields in the noninteracting part of the particle-particle sector of the action must be assigned a different scale dimension than those in the interacting part, and that the latter is logarithmically irrelevant compared to the former. We thus conclude that the Fermi-liquid fixed-point action is given by

$$\mathcal{A}_{\text{FP}} = \frac{1}{V}\sum_{k,\omega}[k + \omega + \gamma\,\omega]\,q^2 + \frac{1}{V}\sum_{k,\omega} P^2, \tag{3.13}$$

where all channels are included, but γ represents the interaction amplitudes in the particle-hole channel only [52].

We thus have shown that there is a choice of scale dimensions that makes the Gaussian part of the action [Eq. (3.7)] a stable fixed-point action, i.e., all other parts of the action are irrelevant with respect to the fixed point. Furthermore, the fixed point describes a Fermi liquid. While the arguments leading to this conclusion are deceptively simple, it is important to realize that their validity relies on two crucial and nontrivial inputs: First, the Ward identity that guarantees the softness of any q vertices (Ref. [19] and Sec. III A), and second, the general loop expansion scheme underlying the renormalization group (Ref. [1]).

We finally mention that, strictly speaking, the Fermi liquid is always unstable with respect to a superconducting state due to the Kohn-Luttinger effect [53]. In the current context, this can be understood as follows. The renormalization process will generate all possible terms quadratic in q, whether or not they are present in the bare action. Physically, these terms are related to to all possible particle-hole and particle-particle interactions in all angular momentum channels. If the net (bare plus RG-generated) interaction in a specific particle-particle

channel is attractive, then the general Gaussian theory yields a superconducting instability at sufficiently low temperatures. Kohn and Luttinger showed that this always happens for sufficiently high angular momenta, and this has also been cast in a RG language [15]. This effect is known to be important only at unrealizably low temperatures, and we neglect it here.

D. Leading corrections to scaling

Now consider the least irrelevant operators in Eqs. (3.10). These are $c_{3,0}$ and $c_{2,1}$ with scale dimensions $-(d-1)/2$, and $c_{4,0}$, $c_{3,1}$, and $c_{2,2}$ with scale dimensions $-(d-1)$. All other terms are more irrelevant by power counting. Furthermore, $c_{3,0}$ and $c_{2,1}$, which multiply odd powers of the fields, enter all observables quadratically, and therefore the generic least irrelevant operator u with respect to the Fermi-liquid fixed point has a scale dimension

$$[u] = -(d-1), \tag{3.14}$$

where u can stand for either of the least irrelevant $c_{n,m}$ or their appropriate squares. We note that $c_{2,2}$ is promoted to the same status as $c_{2,1}^2$, despite the additional P field, by the singular $1/k$ vertex. This is an important difference between clean and disordered electrons. In the latter case there are no singular vertices, and as a result the term of $O(q^2P^2)$ is more irrelevant than the one of $O(q^2P)$ [54]. The same is true of the $\gamma^2\omega^2/k$ vertex in the q^4 term, the analog of which is more irrelevant in a disordered system. As a result, in explicit calculations of leading corrections to Fermi-liquid behavior, there are structurally distinct terms in the clean case that have no analog in the disordered case. Examples will be given in Ref. [35].

The operators collectively denoted by u all become marginal in $d = 1$. This indicates the instability of the Fermi liquid against the formation of a Luttinger-liquid state.

E. Derivation of scaling behavior

We now use the above conclusions to determine the scaling behavior of the observables we are interested in. Let us first consider the DOS. It is given as an expectation value of $\bar{\psi}_n\psi_n$, which is the massive mode P_{nn} introduced in Sec. III B [see Eq. (3.6)]. This couples to the soft mode q via the terms in Eqs. (3.10). Keeping in mind that the P propagator scales as a constant, the DOS can, for scaling purposes, be be expressed as a series of q-correlation functions [55]. Schematically,

$$N \sim 1 + \frac{1}{V^2}\sum_{k,\omega}\langle q^2\rangle + \frac{1}{V^4}\sum_{\{k,\omega\}}\langle q^4\rangle + \dots . \tag{3.15}$$

The RG arguments given above guarantee that the leading contribution to the DOS correction is given by the term quadratic in q. For the scale dimension of the leading scaling part of δN this implies $[\delta N] = 2[q(\boldsymbol{k})] + 2d = 2[q(\boldsymbol{x})] = d - 1$, which in turn implies Eqs. (2.10).

Similarly, the static spin susceptibility is given as a $\langle PP\rangle$ correlation function, and the leading correction to the Fermi-liquid result is given by the terms with coupling constants $c_{2,1}$ and $c_{2,2}$ in Eqs. (3.10). Power counting with the scale dimensions assigned to the fields in Sec. III B shows that its scale dimension is also equal to $d - 1$. Analogously, the leading corrections to the specific-heat coefficient γ_V (which can be expressed as an energy-energy correlation function), and the nematic magnetic susceptibility $\chi_{s,p-\text{wave}}$ are determined by the same terms. We thus have $[\delta\chi_s] = [\delta\chi_{s,p-\text{wave}}] = [\delta\gamma_V] = d - 1$, which yields Eqs. (2.23) and (2.25). Finally, by an analogous argument we find $[\chi_N] = d - 1 - z = d - 2$, which yields Eq. (2.14). We thus have derived scaling from the field theory via a RG treatment.

IV. DISCUSSION AND CONCLUSION

We now discuss various aspects of our approach, and of our results.

A. Alternative scaling analyses

1. *Dependence of observables on the least irrelevant operator*

In Sec. II we assigned a scale dimension to the leading fluctuation corrections to various observables and used those to derive their scaling behavior. Alternatively, one can consider the observables themselves, and use the properties of the Fermi-liquid fixed point, specifically, the scale dimension of the least irrelevant operator with respect to it. This line of reasoning has been used in the past for disordered electrons (see Ref. [54] and Appendix B) and we include it here to show that it is equivalent to the one given in Sec. II E. We illustrate the argument by considering the DOS at zero temperature.

At the Fermi-liquid fixed point the DOS is finite, so we assign it a scale dimension of zero and write, at zero temperature,

$$N(\omega,u) = N(\omega\, b, u\, b^{-(d-1)}). \tag{4.1}$$

Here u is the least irrelevant operator in the action [see Eq. (3.14) and the accompanying discussion]. We now chose $b = 1/\omega$ and obtain

$$N(\omega,u) = N(1, u\,\omega^{(d-1)}) \tag{4.2}$$

Since $N(1,y)$ is evaluated at finite frequency we can Taylor expand in powers of y with impunity and obtain

$$N(\omega \to 0) \propto \text{const} + |\omega|^{(d-1)} + \dots , \tag{4.3}$$

which recovers the result from Sec. II E 1. This line of reasoning is closely related to the one used near critical fixed points to obtain corrections to scaling [1,3].

The same argument can obviously be applied to any other observable. We note, however, that it hinges on the observable under consideration coupling to one of the manifestations of u. If the leading coupling of some quantity were to, say, the coefficient $c_{6,0}$ that is the coupling constant of the q^6 term in Eq. (3.10b), then the leading correction to that quantity would scale as $\omega^{2(d-1)}$, etc.

2. *Corrections to scaling from the fermionic free energy*

An argument related to the one presented in the preceding subsection can be used to obtain all of our results in terms of a single scaling function for the free energy, provided the leading irrelevant variables are taken into account. Consider Eq. (2.4a) again, but take into account the least irrelevant variable u, which reflects bosonic fluctuations. Keeping only

the dependencies on T and H, and in addition on u, we have

$$f^{(f)}(T,H,u) = b^{-2}\, f^{(f)}(Tb, Hb, ub^{-(d-1)}). \tag{4.4}$$

Now if we use the fact that $f^{(f)}(1,0,z)$ and $f^{(f)}(0,1,z)$ are analytic functions of z then we obtain, for example,

$$\gamma_V(T \to 0, H = 0) = \text{const} + c_\gamma\, T^{(d-1)} \tag{4.5a}$$

and

$$\chi_s(T = 0, H \to 0) = \text{const} + c_{\chi_s}\, H^{(d-1)}. \tag{4.5b}$$

Here the c_γ and c_{χ_s} are proportional to u and given in terms of derivatives of the scaling function.

All other results from Sec. II can obviously be obtained by an analogous reasoning. It should be mentioned, however, that this works so easily only since the fermionic and bosonic excitations have the same dynamical scale dimension z. If this is not the case, for instance in the case of disordered electrons, the concept still works but the argument becomes more complicated.

B. Remarks concerning scaling theories

The main point of this paper has been to discuss scaling behavior near the Fermi-liquid fixed point on various levels of sophistication. We add several comments that complement the remarks already made in Ref. [20].

First, scaling works whenever there are soft modes that lead to scale invariance since processes at long wavelengths and low frequencies dominate the relevant physics. Crucial questions are the number and nature of the soft modes, and the observables and external fields they couple to. The concept of scale invariance is best known in the context of critical phenomena, where the relevant soft modes are the critical modes. However, "generic scale invariance," which is caused by soft modes that are due to either Goldstone's theorem or conservation laws, and which holds in entire phases, is at least equally important (see Ref. [4]). The Fermi-liquid phase provides a good example of generic scale invariance, with the soft modes in question the particle-hole excitations, which are Goldstone modes, and the zero-sound and paramagnon collective modes, which are due to conservation laws.

Second, phenomenological scaling relies on input that informs the scaling assumptions. This input may be taken from experiment (as was the case in early studies of critical phenomena), or from theory, which even if incomplete may provide important clues with respect to the above crucial questions, or from both. Even if a complete theory is available, simple scaling is still very useful, as it provides a very simple way to get quick qualitative answers, and to check and elucidate the physics behind explicit calculations.

Third, in a renormalization-group context, scaling near stable fixed points that describe entire phases is just as valid and useful as near critical fixed points. The only difference is the nature of the soft modes (if any; an example of an ordered phase without soft modes is an Ising ferromagnet) that lead to the scale invariance. Moreover, since stable fixed points tend to be characterized by Gaussian fixed-point Hamiltonians, *exact* results can be obtained for physical dimensions [47], which is usually not possible for critical fixed points. Appendix A provides a very simple pedagogical example.

Fourth, since all of the observables we have discussed (except for χ_N) have the same naive dimension, they all generically scale the same way. The only exception is the case of a long-ranged Coulomb interaction, which leads to a second time scale, the plasmon scale, that acts as a dangerous irrelevant variable with respect to the DOS and its susceptibility. This behavior is less generic than, for instance, the case of critical behavior above an upper critical dimension, where dangerous irrelevant variables affect all observables, and naive scaling breaks down. See the next subsection for a discussion of why the DOS is affected by a long-ranged interaction, whereas other observables are not.

C. Gauge invariance and susceptibility of observables to long-range interactions

As we have seen in Sec. II, the nonanalyticities of the DOS and its susceptibility are sensitive to a long-range Coulomb interaction, whereas those of the specific heat and the spin susceptibility are not. This can be understood as follows. For fermions interacting via a Coulomb interaction, the action is invariant under $U(1)$ local gauge transformations, and in particular under a pure imaginary-time transformation $\psi(\boldsymbol{x},\tau) \to \psi(\boldsymbol{x},\tau)\exp[i\Lambda(\tau)]$, $\bar\psi(\boldsymbol{x},\tau) \to \bar\psi(\boldsymbol{x},\tau)\exp[-i\Lambda(\tau)]$ with $\psi(\boldsymbol{x},\tau)$ the fermionic field as a function of position $\boldsymbol{x}$ and imaginary time τ, and $\bar\psi$ the adjoint field. The scalar electromagnetic potential, which is massless and gives rise to the long-range Coulomb interaction, serves as the gauge field. The susceptibilities that determine the specific-heat coefficient and the spin susceptibility are all Fourier transforms of expressions that involve only bilinear products $\bar\psi(\tau)\psi(\tau)$, and hence are gauge invariant. This is not true, however, for the DOS. If we Fourier transform from the imaginary-time variable τ to a Matsubara frequency $\omega_n = 2\pi T(n + 1/2)$ $(n = 0,1,2,\ldots)$, and write $\psi_n \equiv \psi(\omega_n)$, then the DOS is related to a product $Q_{nn} = \bar\psi_n\psi_n$ that is local in Matsubara frequency space rather than in imaginary-time space [see Eqs. (3.2) and (3.6)]. For instance, a linear gauge transformation $\Lambda(\tau) = \alpha\tau$ results in a frequency shift $Q_{nn} \to Q_{n+\alpha,n+\alpha}$. Since the screening of the Coulomb interaction, which results from integrating out the gauge field, is frequency dependent, this makes it plausible that the DOS can be sensitive to the difference between short- and long-ranged interactions. This is also consistent with the fact that the critical behavior at the metal-insulator transition in disordered interaction fermion systems depends on the nature of the interaction for the DOS, but not for gauge invariant quantities [56,57].

D. The DOS anomaly and pseudogaps in two-dimensional electron systems

The linear frequency or energy dependence of the DOS for $d = 2$ [Eqs. (2.10b) and (2.12)] is of interest in the context of the "pseudogap" feature of the DOS that is observed in many strongly correlated $2 - d$ electron systems. This feature, which was first discussed by Mott [58] in the context of generic strongly correlated systems, later became strongly associated with high-T_c superconductivity, and the superconducting gap

in these materials is widely believed to develop out of the pseudogap [59]. However, a recent experiment casts doubt on this notion in at least some superconductors [60]. It is thus possible that at least some of the observed pseudogaps reflect a generic feature of a Fermi liquid, viz., the leading correction to scaling for the DOS, rather than being a harbinger of superconductivity. They still reflect a "strange-metal"-aspect of strongly correlated electrons, however: Since the DOS is the order parameter for the Fermi liquid, strong correlations can induce a quantum phase transition to a non-Fermi-liquid phase where the DOS at the Fermi surface vanishes [17], and the DOS anomaly in the Fermi-liquid phase is a precursor of this transition.

E. Density-of-states fluctuations

To illustrate how strong the effects of the Goldstone modes are on the DOS susceptibility (Sec. II E 2), let us use simple statistical arguments to determine the behavior of $\chi_{\rm OP}$ one would expect in the absence of anomalous fluctuations. Consider $\varphi_n(\boldsymbol{x}) = \psi_n(\boldsymbol{x})/\sqrt{T}$, and $p_n(\boldsymbol{x}) = \bar{\varphi}_n(\boldsymbol{x})\,\varphi_n(\boldsymbol{x})$, and define the "volume" $V_T \equiv 1/T$ in the imaginary-time direction of the space-time of quantum statistical mechanics. Then $\langle p_n(\boldsymbol{x})\rangle \propto V_T$ is a "time-extensive" quantity that is proportional to $V_T \equiv 1/T$. Now consider the fluctuation $\langle(\delta p_n(\boldsymbol{x}))^2\rangle$, the connected part of which is proportional to $1/T = V_T$ (see Appendix E). For the relative fluctuation this implies $\langle(\delta p_n(\boldsymbol{x}))^2\rangle/\langle p_n(\boldsymbol{x})\rangle^2 \propto V_T/V_T^2 = 1/V_T = T$. This just says that the relative fluctuation is proportional to $1/V_T$, as one would expect from ordinary statistics. This yields an estimate for the fluctuations of $\rho(\boldsymbol{x},i\omega_n) = \bar{\psi}_n(\boldsymbol{x})\,\psi_n(\boldsymbol{x})$ which determine $\chi_{\rm N}$ [see Eq. (E1)]:

$$\langle[\delta\rho(\boldsymbol{x},i\omega_n)]^2\rangle \propto \frac{\langle(\delta p_n)^2\rangle}{V_T^2} \propto \frac{\langle(\delta p_n)^2\rangle/V_T^2}{(\langle p_n\rangle/V_T)^2} \propto 1/V_T = T. \tag{4.6}$$

These arguments assume that there are no strong fluctuations in the system that invalidate the simple statistics. For the connected part of $\chi_{\rm N}$, which we denoted by $\chi_{\rm N}^{(1)}$ in Sec. II, we thus have

$$\chi_{\rm N}^{(1)}(\boldsymbol{k},i\omega_n;T) = T\,\theta(\boldsymbol{k},i\omega_n;T). \tag{4.7}$$

In the absence of anomalous fluctuations, θ will scale as the zeroth power of the wave number, the frequency, or the temperature; i.e., $\theta \sim 1$.

We conclude that if the DOS were normally distributed, we would have $\chi_{\rm N}^{(1)} = T \times O(1)$. From Eq. (2.16b) we know that this is not correct even in the case of a short-ranged interaction. Instead, the quantity θ in Eq. (4.7) scales as $\theta \sim 1/k \sim 1/T$ in $d = 2$, and $\chi_{\rm N}^{(1)} \sim 1$. This divergence of the relative DOS fluctuations reflects the strong fluctuations in the system that are a consequence of the existence of the Goldstone modes. In $d = 3$ the fluctuations are weaker, and the dependence of θ on k or T is only logarithmic [see Eq. (2.18b)].

As we have seen in Sec. II, a long-ranged Coulomb interaction further amplifies these effects. Equation (2.21) shows that in $d = 2$, $\theta \sim 1/k^{3/2}$. Even more remarkable is the fact that in both $d = 2$ and $d = 3$, $\chi_{\rm N}$ diverges in the limit of a vanishing wave number $k \to 0$ [see Eqs. (2.21) and (2.22)]. Setting $k = 0$ and considering a finite system with linear dimension L we have, at any nonzero temperature,

$$\chi_N \propto \begin{cases} \ln L & \text{for } d = 2, \\ L & \text{for } d = 3. \end{cases} \tag{4.8}$$

The reason for this unusual behavior is the breakdown of screening of the Coulomb interaction at nonzero frequencies. $\chi_{\rm N}$ is susceptible to both the effects of the Goldstone modes and the breakdown of screening. The effects of the former are stronger in $d = 2$ than in $d = 3$, whereas for the latter the opposite is true. This raises the following interesting point. Consider the quantity $p_n(\boldsymbol{x})$ as defined above, which is subject to thermal and quantum fluctuations that are described by a probability density function P. The expected value $\langle p\rangle = \int D[p]\,p\,P[p]$ exists and determines the density of states. However, the second moment, which determines $\chi_{\rm N}$, does not exist in either $d = 2$ or $d = 3$, and it is easy to see that none of the higher moments exist either. The density of states therefore must have a broad distribution that can not be represented by a Gaussian. This phenomenon requires a separate investigation.

Comparing the results in Sec. II E 2 with the Fermi-liquid result for $\chi_{\rm N}$ we see that the bosonic contributions give the *leading* behavior of $\chi_{\rm N}$ in $d \leqslant 2$. This is in contrast to all other quantities, where the latter give a *correction* to the Fermi-liquid result for all $d > 1$. This is because $\chi_{\rm N}$, as the OP susceptibility, couples particularly strongly to the Goldstone modes. This is precisely analogous to the OP susceptibility in a classical Heisenberg ferromagnet, whose leading behavior in $d \leqslant 4$ is also determined by the coupling to the Goldstone modes (see Appendix A).

We also add some comments about the experimental relevance of the quantity $\chi_{\rm N}$. In any system a local measurement of the DOS depends on the position and is referred to as the local density of states (LDOS). The LDOS gives the dominant contribution to the tunneling current in a scanning tunneling microscope [61]. Its average is the DOS as calculated in this paper and also measured in a tunnel junction. Our OP susceptibility [Eq. (E1)] describes the averaged fluctuations of the LDOS. A suitable two-tip tunneling experiment should be able to give information about this quantity.

F. Conclusion and outlook

In summary, we have presented a scaling analysis of nonanalyticities in Fermi liquids. The most important conclusion is that the exponents of various nonanalyticities that were first derived in perturbation theory are exact. In addition, the scaling theory allows for a unified treatment of clean and disordered electronic systems, as well as various analogous phenomena in classical many-body systems. This demonstrates the generality of the method, which can also be applied to more exotic conductors, such as Dirac and Weyl metals.

ACKNOWLEDGMENTS

We gratefully acknowledge helpful correspondence and discussions with A. Chubukov, S. Gregory, H. Manoharan, and D. Maslov. This work was supported by the National Science Foundation under Grants No. DMR-09-01952 and No.

DMR-09-01907. Part of this work was performed at the Aspen Center for Physics and supported by the NSF under Grant No. PHYS-1066293. We thank the Center for its hospitality.

APPENDIX A: A SIMPLE EXAMPLE: SCALING ANALYSIS OF ϕ^4 THEORY

It is illustrative to recall the scaling analysis for classical ferromagnets that is analogous to our treatment of the clean fermion action in Sec. III. Consider an $O(2)$ ϕ^4 theory with a two-component field $\boldsymbol{\phi}(\boldsymbol{x}) = [\phi_1(\boldsymbol{x}),\phi_2(\boldsymbol{x})]$ and an action

$$S = \int d\boldsymbol{x} \left[\frac{r}{2}\,\boldsymbol{\phi}^2(\boldsymbol{x}) + \frac{c}{2}\,[\nabla\boldsymbol{\phi}(\boldsymbol{x})]^2 + \frac{u}{4}[\boldsymbol{\phi}^2(\boldsymbol{x})]^2 \right]. \quad \text{(A1)}$$

The properties of the ordered phase are usually described by parametrizing $\boldsymbol{\phi}(\boldsymbol{x}) = \rho(\boldsymbol{x})\hat{\boldsymbol{\phi}}(\boldsymbol{x})$, with $\hat{\boldsymbol{\phi}}$ a unit vector [48]. Here we deliberately choose a different parametrization in order to to illustrate our treatment of the fermion problem in the context of a much simpler model.

A saddle-point solution corresponding to the ordered phase is $\boldsymbol{\phi}_{\rm sp}(\boldsymbol{x}) = (\phi_0,0)$ with $\phi_0 = \sqrt{-r/u}$. Now write $\phi_1(\boldsymbol{x}) = \phi_0\,[1 + p(\boldsymbol{x})]$ and $\phi_2(\boldsymbol{x}) = \phi_0\pi(\boldsymbol{x})$, and expand in the fluctuations p and π. Then we obtain a Gaussian action

$$S^{(2)} = \frac{1}{V}\sum_k k^2\,\pi^2 + \frac{1}{V}\sum_k [1 + O(k^2)]\,p^2. \quad \text{(A2)}$$

Here we have performed a Fourier transform from $\pi(\boldsymbol{x})$ and $p(\boldsymbol{x})$ to $\pi(\boldsymbol{k}) \equiv \pi$ and $p(\boldsymbol{k}) \equiv p$, we have rescaled the fields to make the Gaussian coupling constant equal to unity, and we use the same schematic notation as in Sec. III. The non-Gaussian part of the action takes the form

$$\Delta S = \Delta S^{(3)} + \Delta S^{(4)}, \quad \text{(A3)}$$

where

$$\Delta S^{(3)} = \frac{c_{2,1}}{V^2}\sum_{\{k\}}\pi^2\,p + \frac{c_{0,3}}{V^2}\sum_{\{k\}}p^3, \quad \text{(A4a)}$$

$$\Delta S^{(4)} = \frac{c_{4,0}}{V^3}\sum_{\{k\}}\pi^4 + \frac{c_{2,2}}{V^3}\sum_{\{k\}}\pi^2\,p^2 + \frac{c_{0,4}}{V^3}\sum_{\{k\}}p^4. \quad \text{(A4b)}$$

The bare values of the coupling constants $c_{n,m}$ can be expressed in terms of the coupling constants in the original action (A1).

Under renormalization, terms of higher order in the fields are generated, and the coupling constants acquire a wave-number dependence. By symmetry the latter takes the form of a dependence on k^2. Furthermore, the $O(2)$ symmetry of the action leads to a Ward identity that guarantees that the transverse fluctuation π is a soft mode [48]. This is correctly reflected in the Gaussian action (A2). However, ΔS contains terms where π appears without any gradients. The Ward identity ensures that the zeroth-order contributions to these terms in a gradient expansion cancel, and for power-counting purposes $c_{0,4}$, for instance, must be written as

$$c_{4,0} = \tilde{c}_{0,4}k^2, \quad \text{(A5)}$$

and analogously for $c_{2,1}$ and $c_{2,2}$. Indeed, explicitly integrating out p shows that at tree level the term proportional to $c_{2,1}^2$ cancels the term proportional to $c_{4,0}$.

We now assign scale dimensions in an attempt to find a stable fixed point that describes the ferromagnetic phase. We know that π is soft, and that the Goldstone modes are proportional to k^2, which means the first term in Eq. (A2) must be part of the fixed-point action. We also know that the p correlations are short ranged, which implies that the second term in Eq. (A2) is part of the fixed-point action as well. With $[k] = 1$ the scale dimension of the wave number as in Sec. III, this motivates $[\pi] = -(d+2)/2$ and $[p] = -d/2$. Power counting then shows that all terms in Eqs. (A4) are irrelevant. The fixed-point action is thus given by Eq. (A2), and the least irrelevant operator is $\tilde{c}_{4,0}$ with $[\tilde{c}_{4,0}] = -(d-2)$. This implies that $d = 2$ is a lower critical dimension for the problem, consistent with the Mermin-Wagner theorem.

We next add an external magnetic field h in the ϕ_1 direction to the problem. By shifting p one sees that the leading coupling of the field to the soft mode takes the form $h\pi^2$, which gives h a scale dimension $[h] = 2$. Now consider the normalized magnetization $m = \langle\sqrt{1-\pi^2(\boldsymbol{x})}\rangle$, which is the order parameter of the system. The leading fluctuation correction to m is thus given by the correlation function $\delta m = \langle\pi(\boldsymbol{x})\,\pi(\boldsymbol{x})\rangle$, whose scale dimension is $[\delta m] = d-2$. The relevant homogeneity law is thus

$$\delta m(h) = b^{-(d-2)}\delta m(hb^2), \quad \text{(A6)}$$

which yields

$$m \propto \text{const} + h^{(d-2)/2}. \quad \text{(A7)}$$

This nonanalytic field dependence of the magnetization [2] is a result of the Goldstone modes, i.e., the ferromagnons, that are represented by the soft π fluctuations. An equivalent manifestation of the Goldstone modes is the behavior of the longitudinal susceptibility $\chi_{\rm L} = \partial m/\partial h$, which diverges for $h \to 0$ for all $d < 4$ [5]:

$$\chi_{\rm L}(h) \propto h^{-(4-d)/2}. \quad \text{(A8a)}$$

Alternatively, the zero-field inhomogeneous susceptibility diverges for small wave numbers as

$$\chi_{\rm L}(k) \propto k^{-(4-d)}. \quad \text{(A8b)}$$

Note the close analogy between these results for the magnetic order parameter and its susceptibility, and those for the DOS, which is the order parameter for the Fermi-liquid state, and its susceptibility in Sec. II E.

We finally mention that one can integrate out p in a saddle-point approximation that keeps π fixed. For p as defined above Eq. (A2) (before the scaling that normalized the coefficients in the Gaussian action) this leads to

$$p(\boldsymbol{x}) = \sqrt{1-\pi^2(\boldsymbol{x})} - 1 + \frac{c/2u\phi_0^2}{1-\pi^2(\boldsymbol{x})}\nabla^2\sqrt{1-\pi^2(\boldsymbol{x})} + O(\nabla^4). \quad \text{(A9)}$$

Substituting this solution of the saddle-point equation back into the action leads to the familiar nonlinear sigma model

$$S_{\mathrm{NL}\sigma\mathrm{M}} = \frac{c}{2}\phi_0^2 \int d\boldsymbol{x}[[\nabla\pi(\boldsymbol{x})]^2 + (\nabla\sqrt{1-\pi^2(\boldsymbol{x})})^2] + O(\nabla^4). \tag{A10}$$

This derivation of the nonlinear sigma model, which provides an alternative to the usual derivation based on rotational symmetry [48], is the $O(2)$ equivalent of the derivation of an effective action for clean electrons entirely in terms of the soft q field in Ref. [19]. In the electron case, however, the result is *not* a sigma model, and it has not been formulated in a closed form. This is partly due to the more complicated structure of the vertices in the electronic model.

APPENDIX B: SOFT MODES, SCALING, AND NONANALYTICITIES IN DISORDERED ELECTRON SYSTEMS

In this appendix we recall some features of the disordered electron problem, to the extent that they are helpful in understanding the corresponding properties of clean systems discussed in this paper.

The soft-mode effective theory for noninteracting disordered electrons does take the form of a matrix nonlinear sigma model [62] and its generalization to interacting systems adds extra terms to the sigma model [63]. After the analog of the massive field P in Sec. III has been integrated out, the structure of the model is, in the same schematic notation as in Sec. III [54],

$$\mathcal{A}_{\mathrm{NL}\sigma\mathrm{M}} = \frac{1}{V}\sum_{k,\omega}[k^2/G + H\omega + \gamma\omega]q^2 + O(k^2q^4, \omega q^3). \tag{B1}$$

Here q represents the soft components of bilinear fermion fields as in Sec. III, and the nonlinear-sigma-model part of the theory has been expanded in powers of q, keeping only the quadratic term. There are several important differences between this model and the clean model of Ref. [19], despite their apparent similarity. One is that in the disordered case, various observables appear as coupling constants of the field theory. G in Eq. (B1) is proportional to the electrical resistivity, H is proportional to the specific-heat coefficient, and H plus the spin-singlet and spin-triplet interaction constants summarily denoted by γ in Eq. (B1) determine the density and spin susceptibilities, respectively [57]. In contrast, the corresponding observables in the clean case need to be calculated as correlation functions of the basic matrix field. Partly as a result of that, the fixed point describing the disordered Fermi liquid is easier to obtain than in the clean case. Let us assign a scale dimension $[q(\boldsymbol{k})] \equiv [q] = -(d+2)/2$ to the matrix field, and a dynamical exponent $[\omega] = z = 2$ to the frequency. G, H, and γ are then all dimensionless, and the only term shown explicitly in Eq. (B1) represents the fixed-point action. It describes the diffusive modes in a fermion system with quenched disorder. The least irrelevant operators with respect to this fixed point, which we collectively denote by u, all have scale dimensions $[u] = -(d-2)$. This suffices to determine the leading scaling behavior of various observables. For instance, the electrical conductivity σ and the specific-heat coefficient γ_V are both dimensionless according to the above arguments. σ thus obeys a homogeneity law

$$\sigma(\omega,u) = \sigma(\omega b^2, u b^{-(d-2)}) \tag{B2a}$$

which results in a low-frequency nonanalyticity or long-time tail

$$\sigma(\omega \to 0) \propto \mathrm{const} + \omega^{(d-2)/2}. \tag{B2b}$$

Similarly, the specific-heat coefficient obeys

$$\gamma_V(T,u) = \gamma_V(Tb^2, u b^{-(d-2)}) \tag{B3a}$$

which results in

$$\gamma_V(T \to 0) \propto \mathrm{const} + T^{(d-2)/2}. \tag{B3b}$$

The scaling behavior of the DOS and the spin susceptibility can be obtained by analogous arguments. All of these results, which are analogous to the ones for clean systems derived in Secs. III and IV A above, were first derived in perturbation theory [7]. Arguments analogous to those put forward in Sec. III B later showed that they represent the *exact* (as far as the exponents are concerned) leading nonanalyticities [54]. Also note that these scaling arguments immediately show that $d = 2$ is a lower critical dimensionality of the problem, as the disordered Fermi-liquid fixed point becomes unstable for $d \leqslant 2$.

APPENDIX C: LONG-TIME TAILS IN CLASSICAL FLUIDS

Here we sketch how long-time tails in classical fluids can be considered as corrections to scaling at a Navier-Stokes fixed point. Focusing on the viscosity, we first derive the long-time tail as deriving from the dependence of the viscosity on the least irrelevant operator, in analogy to the treatment of clean electrons in Sec. IV A, and of disordered ones in the previous appendix. We then show how, alternatively, it can be understood as the leading scaling behavior of the fluctuation correction to the viscosity, in analogy to the development in Sec. II. Our goal is to demonstrate how universally useful and applicable the notion of corrections to scaling near a stable fixed point is, and that it applies to classical many-body systems as well as to quantum ones.

1. Corrections to scaling at a Navier-Stokes fixed point

We focus on what arguably is the simplest example of a classical long-time tail, viz., the one related to the kinematic viscosity ν. For simplicity, we consider incompressible flow (which in particular eliminates sound waves), and we neglect the pressure-gradient term. The Langevin equation for the transverse fluid velocity $\boldsymbol{u}$ then reads as [64]

$$\partial_t \boldsymbol{u} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} = \nu_0\nabla^2\boldsymbol{u} + \tilde{\boldsymbol{F}}. \tag{C1}$$

Here $\tilde{\boldsymbol{F}}$ is a Gaussian distributed random force whose second moment is fixed by the requirement that Eq. (C1) correctly render the equilibrium velocity fluctuations:

$$\langle \tilde{F}_i(\boldsymbol{x},t)\,\tilde{F}_j(\boldsymbol{x}',t')\rangle \equiv G_{ij}(\boldsymbol{x},t|\boldsymbol{x}',t') = 2T\nu_0\partial_i\partial_j\,\delta(\boldsymbol{x}-\boldsymbol{x}')\,\delta(t-t'). \tag{C2}$$

ν_0 is the bare kinematic viscosity, which gets renormalized to the physical one ν by the nonlinear term in Eq. (C1). We now show how to use renormalization-group and scaling arguments to determine this renormalization. This can be done by using a Martin-Siggia-Rose formalism [65–67] to cast the problem in a field-theoretic language.

We start with the generating functional for all correlation functions of the fluid velocity:

$$Z[\tilde{\boldsymbol{F}}] = \int D[\boldsymbol{u}]\,\delta[\partial_t \boldsymbol{u} + (\boldsymbol{u}\cdot\boldsymbol{\nabla})\boldsymbol{u} - \nu_0\boldsymbol{\nabla}^2\boldsymbol{u} - \tilde{\boldsymbol{F}}] \times e^{-\frac{1}{2}\int d\boldsymbol{x}\,d\boldsymbol{x}'\,dt\,dt'\,\tilde{F}_i(\boldsymbol{x},t)\,G^{-1}_{ij}(\boldsymbol{x},t|\boldsymbol{x}',t')\,\tilde{F}_j(\boldsymbol{x}',t')}. \quad \text{(C3)}$$

Here $D[\boldsymbol{u}]$ is a functional integration measure. Enforcing the functional delta constraint by means of an auxiliary field $\bar{\boldsymbol{u}}$, and integrating out the Langevin noise, we obtain

$$Z \equiv \int D[\tilde{\boldsymbol{F}}]\,Z[\tilde{\boldsymbol{F}}] = \int D[\boldsymbol{u},\bar{\boldsymbol{u}}]\,e^{-S[\boldsymbol{u},\bar{\boldsymbol{u}}]} \quad \text{(C4a)}$$

with an action

$$S[\boldsymbol{u},\bar{\boldsymbol{u}}] = i\int d\boldsymbol{x}\,dt\,\bar{\boldsymbol{u}}\cdot[\partial_t\boldsymbol{u} + c\,(\boldsymbol{u}\cdot\boldsymbol{\nabla})\boldsymbol{u} - \nu_0\boldsymbol{\nabla}^2\boldsymbol{u}] + \frac{1}{2}\int d\boldsymbol{x}\,\boldsymbol{x}'\,dt\,dt'\,\bar{u}_i(\boldsymbol{x},t)\,G_{ij}(\boldsymbol{x},t|\boldsymbol{x}',t')\,\bar{u}_j(\boldsymbol{x}',t'). \quad \text{(C4b)}$$

Here we have introduced a nominal coupling constant c for the nonlinear term, whose bare value is $c_0 = 1$.

Now we assign scale dimensions $[L] = -1$ and $[t] = -z$ to length and time, respectively. Choosing

$$[\boldsymbol{u}] = [\bar{\boldsymbol{u}}] = d/2 \quad \text{(C5a)}$$

and

$$z = 2 \quad \text{(C5b)}$$

leads to a stable Navier-Stokes fixed point with respect to which the nonlinear term is irrelevant. Indeed, the scale dimension of the coupling constant c is negative for $d > 2$,

$$[c] = -(d-2)/2, \quad \text{(C6)}$$

while ν_0 is dimensionless and the viscosity term is thus part of the fixed-point action. Since the nonlinear term is cubic in the fields, it always appears squared in explicit calculations. The wave-number and frequency-dependent kinematic viscosity thus obeys a homogeneity law

$$\nu(k,\omega) = f_\nu(kb,\omega b^2,c^2 b^{-(d-2)}) \quad \text{(C7)}$$

with f_ν a scaling function. From Eq. (C7) we obtain, in particular, by expanding in powers of the small third argument,

$$\nu(k=0,\omega) \propto 1 + \text{const}\times\omega^{(d-2)/2}, \quad \text{(C8)}$$

where the constant is proportional to c^2. This is the well-known classical long-time tail [12]. In $d = 2$, c becomes marginal, which reflects the fact that the local description of hydrodynamics breaks down in $d \leqslant 2$ [12,68].

2. Scaling of the fluctuation correction

To complete the analogy with our various discussions of the quantum problem, we now consider the long-time tail of the classical viscosity from an alternative point of view. The kinematic viscosity ν is defined as the shear viscosity η divided by the mass density ρ, $\nu = \eta/\rho$, and thus has a naive dimension of a length squared divided by a time. With the choice of scale dimensions specified in Sec. C 1, this makes ν dimensionless, consistent with the fixed-point action identified above. Now consider the leading correction $\delta\nu$ to the kinematic viscosity. To determine the scale dimension of $\delta\nu$, we recall that the viscosity physically results from a frictional force $\boldsymbol{F}_f$. With G a friction coefficient, we write

$$\boldsymbol{F}_f = G\boldsymbol{u}. \quad \text{(C9)}$$

Considering planar Couette flow within a hypercube of linear dimension L we have, for the x component F_f of $\boldsymbol{F}_f$, and with an accuracy that suffices for dimensional arguments,

$$F_f = GL\partial u_x/\partial y. \quad \text{(C10)}$$

On the other hand, the x-y component of the stress tensor T is given by

$$T_{xy} = \eta\partial u_x/\partial y. \quad \text{(C11)}$$

To obtain the frictional force we need to multiply by the cross-sectional area L^{d-1}. Equating the result with Eq. (C10) we obtain

$$GL\partial u_x/\partial y = L^{d-1}\eta\,\partial u_x/\partial y, \quad \text{(C12)}$$

and finally

$$G = L^{d-2}\eta. \quad \text{(C13)}$$

G/ρ thus scales as L^{d-2} times a dimensionless quantity, which yields $[\delta\nu] = -(d-2)$ for the scale dimension of $\delta\nu$. The appropriate homogeneity law, according to Eq. (2.1), is thus

$$\delta\nu(k,\omega) = b^{(d-2)}\delta\nu(kb,\omega b^2). \quad \text{(C14)}$$

Equation (C8) now follows as the leading scaling behavior of $\delta\nu$.

The above arguments, which are physically equivalent to the ones given in Appendix C 1, are analogous to the scaling theory of electron localization by Abrahams *et al.* [69]. Building on arguments by Thouless, these authors realized that the natural scaling variable is the conductance G, which is related to the conductivity σ by $G = L^{d-2}\sigma$. The relation between the friction coefficient G and the viscosity η in Eq. (C13) is the precise analog for the classic fluid case.

APPENDIX D: DENSITY OF STATES, AND ITS SUSCEPTIBILITY, IN THE LONG-RANGE CASE

Here we explain the origin of the results in the long-range case that were given in Sec. II E. For the case of a short-range interaction, the expression for the DOS [Eqs. (3.6)] were evaluated in Ref. [19] to one-loop order. For the causal function $Q(i\omega_n)$ [Eq. (3.6b)], with $\omega_n > 0$ in d dimensions, it takes the

form [see Eq. (5.4) in that reference]

$$Q(i\omega_n) = 2i\,\frac{1}{V}\sum_{\boldsymbol{k}}\frac{1}{v_{\rm F}^2k^2}\,T\sum_{m<0}\varphi_d'(i\Omega_{n-m}/v_{\rm F}k) \times\frac{2N_{\rm F}\gamma}{1-2N_{\rm F}\gamma\,\varphi_d(i\Omega_{n-m}/v_{\rm F}k)\Omega_{n-m}/v_{\rm F}k}. \quad \text{(D1a)}$$

Here γ is the interaction amplitude, and

$$\varphi_d(z) = \frac{i}{z}\,{}_2F_1(1,1/2,d/2;1/z^2) \quad \text{(D1b)}$$

with ${}_2F_1$ Gauss's hypergeometric function, and $\varphi_d'(z) = d\varphi_d(z)/dz$. Evaluating this integral, and using Eq. (3.6a), yields results that are consistent with Eqs. (2.10).

In the case of a long-ranged Coulomb interaction, the amplitude γ gets replaced by a wave-number-dependent function that has the form of a statically screened Coulomb potential:

$$\gamma \to \gamma(k) = \frac{1}{2N_{\rm F}}\,\frac{\kappa^{d-1}}{k^{d-1}+\kappa^{d-1}}, \quad \text{(D2)}$$

where κ is the screening wave number. At zero temperature, the integral can be written as

$$Q(i\omega) = Q(i0) - \frac{S_{d-1}}{\pi(2\pi)^d v_{\rm F}^d}\,\omega^{d-1}\int_0^{v_{\rm F}\Lambda/\omega} dy\, y^{d-2} \times\int_0^{1/y} dx\,\frac{i\varphi_d'(ix)}{1+(\omega/v_{\rm F}\kappa)^{d-1}y^{d-1}-x\varphi_d(ix)}. \quad \text{(D3)}$$

Asymptotic analysis reveals that the function of ω given by the double integral on the right-hand side has, for $\omega\to 0$, a constant contribution and a contribution proportional to $\omega^{(d-2)/(d-3)}$. For $d<2$, the latter is dominant and leads to Eq. (2.12).

An analogous, albeit more involved, calculation, the structure of which was given in Ref. [39], yields the behavior of the DOS susceptibility given in Sec. II E 2.

APPENDIX E: STRUCTURE OF THE DENSITY-OF-STATES SUSCEPTIBILITY

The density-of-states susceptibility $\chi_{\rm N}$ is given as a four-fermion correlation function [39]

$$\chi_{\rm N}(\boldsymbol{x}-\boldsymbol{y};i\omega_n,i\omega_m) = \langle\delta\rho(\boldsymbol{x},i\omega_n)\,\delta\rho(\boldsymbol{y},i\omega_m)\rangle. \quad \text{(E1)}$$

Here $\rho(\boldsymbol{x},i\omega_n) = \bar\psi_n(\boldsymbol{x})\,\psi_n(\boldsymbol{x})$ and $\delta\rho = \rho - \langle\rho\rangle$. Defining a Fourier transform from a Matsubara frequency ω_n to an imaginary time τ as in Ref. [39], this takes the form

$$\chi_{\rm N}(\boldsymbol{x}-\boldsymbol{y};i\omega_n,i\omega_m) = T^2\int_0^{1/T} d\tau_1\,d\tau_2\,d\tau_3\,d\tau_4\;e^{i\omega_n(\tau_1-\tau_2)+i\omega_m(\tau_3-\tau_4)} \times\langle\delta[\bar\psi(\boldsymbol{x},\tau_1)\,\psi(\boldsymbol{x},\tau_2)]\,\delta[\bar\psi(\boldsymbol{y},\tau_3)\,\psi(\boldsymbol{y},\tau_4)]\rangle. \quad \text{(E2)}$$

There are two distinct contributions to this correlation function: First, a disconnected one in which the four-fermion correlation factorizes into a product of two two-fermion correlations:

$$\chi_{\rm N}^{\rm dc}(\boldsymbol{x}-\boldsymbol{y};i\omega_n,i\omega_m) = -T^2\int_0^{1/T} d\tau_1\,d\tau_2\,d\tau_3\,d\tau_4\;e^{i\omega_n(\tau_1-\tau_2)+i\omega_m(\tau_3-\tau_4)} \times\langle\bar\psi(\boldsymbol{x},\tau_1)\psi(\boldsymbol{y},\tau_4)\rangle\langle\bar\psi(\boldsymbol{y},\tau_3)\psi(\boldsymbol{x},\tau_2)\rangle, \quad \text{(E3a)}$$

and second, a connected one that contains the contributions to $\langle\bar\psi\psi\bar\psi\psi\rangle$ that do not factorize:

$$\chi_{\rm N}^{\rm c}(\boldsymbol{x}-\boldsymbol{y};i\omega_n,i\omega_m) = T^2\int_0^{1/T} d\tau_1\,d\tau_2\,d\tau_3\,d\tau_4\;e^{i\omega_n(\tau_1-\tau_2)+i\omega_m(\tau_3-\tau_4)} \times\langle\bar\psi(\boldsymbol{x},\tau_1)\psi(\boldsymbol{x},\tau_2)\bar\psi(\boldsymbol{y},\tau_3)\psi(\boldsymbol{y},\tau_4)\rangle^{\rm c}. \quad \text{(E3b)}$$

Using time translational invariance, we finally obtain

$$\chi_{\rm N}^{\rm dc}(\boldsymbol{x}-\boldsymbol{y};i\omega_n,i\omega_m) = -\delta_{nm}\int_0^{1/T} d\tau\,d\tau'\,\mathcal{G}(\boldsymbol{x}-\boldsymbol{y},\tau)\,\mathcal{G}(\boldsymbol{y}-\boldsymbol{x},\tau'), \quad \text{(E4a)}$$

where $\mathcal{G}(\boldsymbol{x},\tau) = \langle\bar\psi(\boldsymbol{x},\tau)\,\psi(0,0)\rangle$, and

$$\chi_{\rm N}^{\rm c}(\boldsymbol{x}-\boldsymbol{y};i\omega_n,i\omega_m) = T\int_0^{1/T} d\tau_1\,d\tau_2\,d\tau_3\;e^{i\omega_n(\tau_1-\tau_2)+i\omega_m(\tau_3)} \times\langle\bar\psi(\boldsymbol{x},\tau_1)\psi(\boldsymbol{x},\tau_2)\bar\psi(\boldsymbol{y},\tau_3)\psi(\boldsymbol{y},0)\rangle^{\rm c}. \quad \text{(E4b)}$$

This shows that $\chi_{\rm N}^{\rm dc}$ approaches a constant as $T\to 0$, whereas $\chi_{\rm N}^{\rm c}$ is proportional to T. These are the properties we used in Sec. II, where we denoted $\chi_{\rm N}^{\rm dc}$ and $\chi_{\rm N}^{\rm c}$ by $\chi_{\rm N}^{(0)}$ and $\chi_{\rm N}^{(1)}$, respectively.

[1] K. G. Wilson and J. Kogut, Phys. Rep. **12**, 75 (1974).

[2] S.-K. Ma, *Modern Theory of Critical Phenomena* (Benjamin, Reading, MA, 1976).

[3] J. Cardy, *Scaling and Renormalization in Statistical Physics* (Cambridge University Press, Cambridge, 1996).

[4] D. Belitz, T. R. Kirkpatrick, and T. Vojta, Rev. Mod. Phys. **77**, 579 (2005).

[5] E. Brézin and D. J. Wallace, Phys. Rev. B **7**, 1967 (1973).

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

[7] B. L. Altshuler and A. G. Aronov, in *Electron-Electron Interactions in Disordered Systems*, edited by M. Pollak and A. L. Efros (North-Holland, Amsterdam, 1984).

[8] At zero temperature there is a unique ground state with a unique energy, and what is conventionally called the density of states is related to the spectrum of the single-particle Green function [see Eq. (3.2) later in this paper]. To calculate observables at any nonzero temperature, in either a canonical or a grand canonical ensemble, one averages over all energies with the appropriate Boltzmann weight. The spectrum of the finite-temperature

Green function is then related to a temperature-dependent, or thermal, density of states. The density of states thus becomes a thermal quantity because of the statistical-mechanics average over all states. The quantity measured in a tunneling experiment is this thermal density of states.

[9] B. J. Alder and T. E. Wainwright, Phys. Rev. A **1**, 18 (1970).

[10] M. H. Ernst, E. H. Hauge, and J. M. J. van Leeuwen, Phys. Rev. Lett. **25**, 1254 (1970).

[11] J. R. Dorfman and E. G. D. Cohen, Phys. Rev. Lett. **25**, 1257 (1970).

[12] Y. Pomeau and P. Resibois, Phys. Rep. **2**, 63 (1975).

[13] J. R. Dorfman, T. R. Kirkpatrick, and J. V. Sengers, Annu. Rev. Phys. Chem. **45**, 213 (1994).

[14] G. Baym and C. Pethick, *Landau Fermi-Liquid Theory* (Wiley, New York, 1991).

[15] R. Shankar, Rev. Mod. Phys. **66**, 129 (1994).

[16] T. Giamarchi, *Quantum Physics in One Dimension* (Clarendon, Oxford, 2004).

[17] T. R. Kirkpatrick and D. Belitz, Phys. Rev. Lett. **108**, 086404 (2012).

[18] W. Metzner, M. Salmhofer, C. Honerkamp, V. Meden, and K. Schönhammer, Rev. Mod. Phys. **84**, 299 (2012).

[19] D. Belitz and T. R. Kirkpatrick, Phys. Rev. B **85**, 125126 (2012).

[20] The term 'scaling theory' is associated with a variety of meanings. At one level, it refers to the phenomenological considerations first applied by Widom, Kadanoff, Fisher, and others to describe critical phenomena (see, e.g., Refs. [70, 71]. It relies on unproven assumptions, which are usually informed by either experimental observations, or insight from microscopic theories (sometimes incomplete ones), or both, and by itself has no microscopic basis. However, the fact that it requires, in its simplest form, nothing but dimensional analysis makes it a very useful and powerful tool. At a different level, the term refers to the analysis of an effective action or Hamiltonian by means of RG techniques, which can be used to derive the assumptions the phenomenological scaling theory is based on. These derivations can take various rather different forms; the version we use is due to Ma [2]. In this paper we use phenomenological scaling arguments in Sec. II, and RG-based scaling in Sec. III We also note that we consider scaling at a stable RG fixed point that describes an ordered phase, namely, the Fermi-liquid phase, that is analogous to the stable fixed point that describes the ordered phase in ϕ^4 theory, as opposed to a critical fixed point that describes a continuous phase transition. See Appendix A, Sec. IV B, and the book by Ma, Ref. [2], for additional discussions of these points.

[21] We write $a \sim b$ for "a scales as b," $a \propto b$ for "a is proportional to b," and $a \cong b$ for "a is isomorphic to b."

[22] For instance, the density of states N is an inverse energy times an inverse volume, so $[N] = -[E] - [V] = d - z$, etc. There is some ambiguity here, since one can choose a different value for $[A]$, which is compensated by the scale dimension of some variable A depends on. However, using the naive scale dimension is arguably the easiest and most transparent choice, and we adopt this convention.

[23] D. Pines and P. Nozières, *The Theory of Quantum Liquids* (Addison-Wesley, Redwood City, CA, 1989).

[24] T. R. Kirkpatrick and D. Belitz, Phys. Rev. Lett. **106**, 105701 (2011).

[25] T. R. Kirkpatrick and D. Belitz, Phys. Rev. B **85**, 134451 (2012).

[26] This structure is simpler than the case of electrons in the presence of quenched disorder, which has more universality classes [56, 57, 72]. In the disordered case, magnetic impurities and spin-orbit scattering provide additional universality classes. The former has no analog in the clean case, and the spin-orbit interaction in the clean case just leads to an effective interaction in the particle-hole spin-triplet channel that does not change the soft-mode structure.

[27] How many scaling parts there are in total is *a priori* not clear. Classes of soft modes, and associated scaling parts, in addition to the two considered here may exist; if they are weakly coupled they will all appear additively for asymptotic scaling purposes. However, there is substantial evidence that the bosonic two-particle excitations provide the leading corrections to scaling at the Fermi-liquid fixed point. In addition to the RG analysis of a formally exact effective field theory provided in Sec. III, this evidence consists of the consistency with many different explicit perturbative calculations (see Ref. [35]).

[28] L. D. Landau and E. M. Lifshitz, *Statistical Physics Part 1* (Butterworth Heinemann, Oxford, 1980).

[29] D. Belitz, T. R. Kirkpatrick, and T. Vojta, Phys. Rev. B **55**, 9452 (1997).

[30] G. Y. Chitov and D. Sénéchal, Phys. Rev. B **52**, 13487 (1995).

[31] C. Nayak and F. Wilczek, Int. J. Mod. Phys. B **10**, 847 (1996).

[32] N. Dupuis and G. Y. Chitov, Phys. Rev. B **54**, 3040 (1996).

[33] G. Y. Chitov and D. Sénéchal, Phys. Rev. B **57**, 1444 (1998).

[34] $N(\omega)$ is the spectrum of a causal function that is an odd function of the complex frequency (see Sec. III A). As a result, N is an even function that depends on $|\omega|$.

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

[36] E. G. Mishchenko and A. V. Andreev, Phys. Rev. B **65**, 235310 (2002).

[37] A. V. Chubukov, D. L. Maslov, S. Gangadharaiah, and L. I. Glazman, Phys. Rev. B **71**, 205112 (2005).

[38] Whether at two-loop or higher order an $|\omega|$ behavior with a nonzero prefactor emerges is an open question.

[39] T. R. Kirkpatrick and D. Belitz, Europhys. Lett. **102**, 17002 (2013).

[40] J. Betouras, D. Efremov, and A. Chubukov, Phys. Rev. B **72**, 115112 (2005).

[41] G. Y. Chitov and A. J. Millis, Phys. Rev. B **64**, 054414 (2001).

[42] A. V. Chubukov and D. L. Maslov, Phys. Rev. B **68**, 155113 (2003).

[43] D. L. Maslov, A. V. Chubukov, and R. Saha, Phys. Rev. B **74**, 220402(R) (2006).

[44] D. Belitz, T. R. Kirkpatrick, and T. Vojta, Phys. Rev. Lett. **82**, 4707 (1999).

[45] T. R. Kirkpatrick and D. Belitz, Phys. Rev. B **53**, 14364 (1996).

[46] D. V. Khveshchenko and M. Reizer, Phys. Rev. B **57**, R4245 (1998).

[47] This statement assumes that the scale dimensions of all of the pertinent quantities, in particular the field conjugate to the order parameter, can be simply obtained. This is not necessarily the case for all Gaussian theories. A counterexample is the fermionic formulation of the two-dimensional classical Ising model [73, 74], where the fixed-point action is Gaussian, yet the critical exponents can not be obtained simply by power counting. We would like to thank an anonymous referee for pointing this out.

[48] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena* (Oxford University Press, Oxford, 1996).

[49] We use what in Ref. [19] was called the density formulation of the effective theory, within which q and P are local in space.
[50] The theory also contains a second pair of fields λ and Λ, which technically arise from constraining bilinear products of fermionic fields to the bosonic fields q and P according to Eq. (3.1). Their propagators are minus the noninteracting parts of the q and P propagators, respectively. In the non-Gaussian part of the action they always appear in the combinations $\not{q} = q - \lambda$ and $\not{P} = P - \Lambda$, whose propagators are thus the q and P propagators with the noninteracting parts subtracted. See Ref. [19] for details.
[51] P. A. Lee and T. V. Ramakrishnan, Rev. Mod. Phys. **57**, 287 (1985).
[52] This conclusion is not consistent with the treatment of the Cooper channel in several recent papers [75–77]. However, the discrepancies, which can be traced to the role of the Cooper ladder resummation in a RG scheme, are only logarithmic and can be suppressed by considering a system in a weak magnetic field, as mentioned in the text. We therefore do not dwell on this point here and will revisit it in a future publication [35] where we will discuss the Cooper channel more explicitly.
[53] W. Kohn and J. M. Luttinger, Phys. Rev. Lett. **15**, 524 (1965).
[54] D. Belitz and T. R. Kirkpatrick, Phys. Rev. B **56**, 6513 (1997).
[55] This is in analogy to the fact that the massive field σ in a classical $O(N)$ nonlinear sigma model can be expressed as a power series in the soft field $\boldsymbol{\pi}$, see. e.g., Ref. [48].
[56] A. M. Finkelstein, Zh. Eksp. Teor. Fiz. **86**, 367 (1984) [Sov. Phys. JETP **59**, 212 (1984)].
[57] D. Belitz and T. R. Kirkpatrick, Rev. Mod. Phys. **66**, 261 (1994).
[58] N. F. Mott, Rev. Mod. Phys. **40**, 677 (1968).
[59] T. Timusk and B. Statt, Rep. Prog. Phys. **62**, 61 (1999).
[60] T. Jacobs, S. O. Katterwe, H. Motzkau, A. Rydh, A. Maljuk, T. Helm, C. Putzke, E. Kampert, M. V. Kartsovnik, and V. M. Krasnow, Phys. Rev. B **86**, 214506 (2012).
[61] J. Tersoff and D. R. Hamann, Phys. Rev. B **31**, 805 (1985).
[62] F. Wegner, Z. Phys. B **35**, 207 (1979).
[63] A. M. Finkelstein, Zh. Eksp. Teor. Fiz. **84**, 168 (1983) [Sov. Phys. JETP **57**, 97 (1983)].
[64] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon, Oxford, 1987).
[65] P. C. Martin, E. D. Siggia, and H. A. Rose, Phys. Rev. A **8**, 423 (1973).
[66] R. Bausch, H. K. Janssen, and H. Wagner, Z. Phys. B **24**, 113 (1976).
[67] C. DeDominicis and L. Peliti, Phys. Rev. B **18**, 353 (1978).
[68] D. Forster, D. R. Nelson, and M. J. Stephen, Phys. Rev. A **16**, 732 (1977).
[69] E. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan, Phys. Rev. Lett. **42**, 673 (1979).
[70] L. P. Kadanoff, W. Götze, D. Hamblen, R. Hecht, E. A. S. Lewis, V. V. Palciauskas, M. Rayl, J. Swift, D. Apsens, and J. Kane, Rev. Mod. Phys. **39**, 395 (1967).
[71] E. Stanley, *Introduction to Phase Transitions and Critical Phenomena* (Oxford University Press, Oxford, 1971).
[72] K. B. Efetov, A. I. Larkin, and D. E. Khmelnitskii, Zh. Eksp. Teor. Fiz. **79**, 1120 (1980) [Sov. Phys.–JETP **52**, 568 (1980)].
[73] T. D. Schultz, D. C. Mattis, and E. H. Lieb, Rev. Mod. Phys. **36**, 856 (1964).
[74] V. N. Plechko, J. Phys. Studies **3**, 312 (1999).
[75] A. Shekhter and A. M. Finkelstein, Proc. Natl. Acad. Sci. USA **103**, 15765 (2006).
[76] A. V. Chubukov and D. L. Maslov, Phys. Rev. B **76**, 165111 (2007).
[77] H. Meier, C. Pépin, and K. B. Efetov, Phys. Rev. B **84**, 205131 (2011).