Effective soft-mode theory of strongly interacting fermions

D. Belitz and T. R. Kirkpatrick

Phys. Rev. B **85**, 125126 — Published 23 March 2012

DOI: 10.1103/PhysRevB.85.125126

Effective Soft-Mode Theory of Strongly Interacting Fermions

D. Belitz[1,2], T. R. Kirkpatrick[3]
[1] *Department of Physics and Institute of Theoretical Science, University of Oregon, Eugene, OR 97403, USA*
[2] *Materials Science Institute, University of Oregon, Eugene, OR 97403, USA*
[3] *Institute for Physical Science and Technology, and Department of Physics, University of Maryland, College Park, MD 20742, USA*

An effective field theory for clean electron systems is developed in analogy to the generalized nonlinear sigma model for disordered interacting electrons. The physical goal is to separate the soft or massless electronic degrees of freedom from the massive ones and integrate out the latter to obtain a field theory in terms of the soft degrees of freedom only. The resulting theory is not perturbative with respect to the electron-electron interaction. It is controlled by means of a systematic loop expansion and allows for a renormalization-group analysis in a natural way. It is applicable to universal phenomena within phases, and to transitions between phases, with order parameters in arbitrary angular-momentum channels, and in the spin-singlet, spin-triplet, particle-hole, and particle-particle channels. Applications include ferromagnetic and ferrimagnetic ordering, non-s-wave ferromagnetic order (magnetic nematics), Fermi-liquid to non-Fermi-liquid transitions, and universal phenomena within a Fermi-liquid phase.

PACS numbers: 71.10.-w; 71.10.Pm; 71.10.Hf; 71.10.+a

I. INTRODUCTION

The theory of many-fermion systems is of obvious importance because of its ubiquitous applications in condensed-matter physics. It is also a very hard problem, due to the difficulty of dealing with a macroscopic number of electrons which, in many materials, interact strongly enough to preclude treating the interactions as weak perturbations of noninteracting electrons. Historically, there have been two main approaches to the problem. Landau's Fermi-liquid theory,[1,2] although not perturbative, is phenomenological and self-contained in nature, and often a more microscopic approach is desirable, especially for investigations of the limitation of Fermi-liquid theory. Many-body diagrammatic techniques,[3,4] on the other hand, are limited by their inherently perturbative approach. Furthermore, it is neither feasible nor desirable to study all aspects of a microscopic model, except in some special cases. For many important applications it suffices to keep only the low-lying excitations or soft modes of the system, which govern the behavior at long times and large distances. The most obvious examples of such applications are phase transitions, but there also are many qualitative properties of entire phases that depend on the soft modes only. These observables also have the attractive property of being universal in the sense that they depend only on basic symmetry properties of the system, and not on any microscopic details. Consider, for instance, a Heisenberg ferromagnet. The critical behavior at the Curie point, i.e., the phase transition from the paramagnetic to the ferromagnetic phase, is entirely governed by soft modes,[5] but so is the divergence of the longitudinal susceptibility everywhere in the ordered phase.[6] To describe and understand either phenomenon it therefore suffices to consider an effective theory that takes into account only the relevant soft modes. In the case of the Heisenberg ferromagnet at temperatures $T > 0$, either a classical $O(3)$ ϕ^4-theory or an $O(3)$ nonlinear sigma model provides a suitable effective theory.[7] Such a theory can sometimes be constructed based on symmetry considerations only, but it often is desirable, and leads to greater physical insight, to start with a microscopic description and derive the effective theory by integrating out all non-essential degrees of freedom in some simple approximation that respects the crucial symmetries. The resulting "hydrodynamic", i.e., long-wavelength and low-frequency, description of any system of interest is complementary to, and serves a very different purpose than, for instance, a first-principles approach that aims at a description at atomic length and energy scales. In order to obtain a well-behaved effective theory, care must be taken to keep all soft modes of interest *and* all other soft modes that couple to them. Integrating out soft modes leads to long-ranged effective interactions between the remaining degrees of freedom, and hence to non-local effective field theories that can be hard to handle technically.

For noninteracting disordered electrons such an effective-theory approach was pioneered by Wegner.[8] Guided by an analogy with classical ferromagnets he constructed a matrix nonlinear sigma model capable of describing the instability of a (disordered) Fermi-liquid phase against the formation of an Anderson insulator, and allowed for the determination of the critical behavior at the Anderson transition in terms of an ϵ-expansion about the lower critical dimension for this problem, $d_c^- = 2$. Wegner's original formulation was expanded upon by various authors,[9–11] and eventually

his sigma model was generalized to interacting disordered electrons.[12] This theory and its generalizations have been extensively used to describe the Anderson-Mott metal-insulator transition.[13,14] These theories are fundamentally based on diffusive soft modes, and do not allow for the clean limit to be taken. This led to the strange situation where a successful effective soft-mode theory was available for disordered electrons, but not for clean ones, although one might think that the latter poses a simpler problem.[15] This last conclusion is fallacious, however, mostly because the number of soft modes is much larger in clean electron systems than in disordered ones. This was noted in Ref. 16, whose main objective was to provide a thorough derivation of Finkelstein's generalized nonlinear sigma model for disordered systems. While this theory did allow for the clean limit to be taken, the resulting field theory had non-local vertices and did not offer any obvious way to obtain a description as useful as the nonlinear sigma model for the disordered case. Another effective field theory for clean electrons was developed in Ref. 17. This theory is technically based on a supersymmetric matrix formulation and aims to incorporate and unify various other approaches that discuss the validity and limitations of Landau Fermi-liquid theory.[18–23] It introduces bosonic degrees of freedom and integrates out the fermions, as did Ref. 16, in contrast to Shankar's derivation of Fermi-liquid theory that applied renormalization-group (RG) techniques to a fermionic field theory.[24] The theory of Ref. 17 has recently been refined[25] and formulated both in a version that is formally exact and may be better suitable for numerical work than standard Monte-Carlo techniques,[26] and as an effective field theory for low-lying excitations.[27] The latter was used for an analysis of the low-temperature specific heat in two-dimensional systems.

In the present paper we take a different approach to this problem. It is similar in spirit to Refs. 17,27, but rather different technically. Our physical motivation is the desirability of a broadly useful effective theory that is capable of describing widely different phenomena such as ferromagnetic ordering in non-s-wave angular momentum channels,[28,29] quantum ferrimagnetic order, which has received little attention so far, a possible breakdown of Fermi-liquid theory due to a vanishing quasi-particle weight,[30] and, in addition, novel universal aspects of the Fermi-liquid phase.[30] We are interested in a local field theory that is not perturbative with respect to the electron-electron interaction, but rather can be controlled by means of a systematic loop expansion. This will provide a basis for an application of RG techniques that allow for a resummation of the expansion in powers of the loop parameter, and thus allow to go beyond perturbation theory in a controlled fashion. This not possible within the framework of traditional perturbation theory,[3,4] which provides no information about the structure of the renormalized theory. Technically, the theory we will develop is based on a generalization of Ref. 16. The reason for the non-locality of the latter was its restriction to density or s-wave ($\ell = 0$) modes, with soft modes in higher angular-momentum channels being effectively integrated out. Formulating the theory in terms of phase-space variables allows one to keep the soft modes in all angular-momentum channels. This leads both to a local theory and allows for natural description of order in $\ell \neq 0$ channels. Integrating out the massive modes in a tree approximation leads to an effective action that explicitly keeps all of the soft modes. Its structure is different from that of the generalized nonlinear sigma model for the disordered case. It can be constructed to any desired order in the soft degrees of freedom, which in turn allows to perform a loop expansion to any desired order.

The organization of this paper is as follows. In Sec. II we formulate our model in terms of fermionic fields. We then bosonize the theory by constraining bilinear products of fermion fields to bosonic matrix fields Q and integrating out the fermions. We then discuss symmetry properties of the Q-fields and representations of observables in terms of Q-correlation functions. In Sec. III we identify the soft modes of the action by means of a Ward identity. In Sec. IV we expand the Q-field theory about a saddle-point solution that describes a Fermi liquid, and separate the soft fluctuations from the massive ones by means of the Ward identity. The massive degrees of freedom are then integrated out in a tree approximation to derive an effective theory entirely in terms of the soft modes. We also make contact with the density formulation of Ref. 16. In Sec. V we discuss universal features of the energy-dependent density of states (DOS) in a Fermi-liquid as a simple application. In Sec. VI we summarize and discuss the effective theory. We finally discuss various future applications of the theory. A discussion of the effects of quenched disorder, and a pedagogical discussion of the $O(2)$ nonlinear sigma model that stresses analogies with the current theory, are relegated to two appendixes.

II. FIELD-THEORETIC FORMULATION OF THE INTERACTING FERMION PROBLEM

A. Fermionic formulation

Our starting point is a description of interacting fermions in terms of Grassmann fields. The partition function can be written[31]

$$Z = \int D[\bar{\psi}, \psi]\; e^{S[\bar{\psi},\psi]} \quad . \tag{1}$$

Here $\bar{\psi}$ and ψ are Grassmann-valued fields , and $D[\bar{\psi},\psi]$ is the Grassmannian integration measure. The action S has the form

$$S = \int dx \sum_\sigma \bar{\psi}_\sigma(x)\ [-\partial_\tau - \epsilon(\boldsymbol{\nabla}) + \mu]\ \psi_\sigma(x) + S_{\rm int}\ . \tag{2a}$$

The first term describes noninteracting electrons with chemical potential μ. For simplicity and definiteness we will assume a single parabolic band, i.e.,

$$\epsilon(\boldsymbol{\nabla}) = -\boldsymbol{\nabla}^2/2m_{\rm e} \tag{2b}$$

with $m_{\rm e}$ the effective electron mass. If desired, the model can easily be generalized to include a nontrivial band structure instead. We use a $(d+1)$-vector space-time notation with $x = (\boldsymbol{x},\tau)$, $\int dx = \int_V d\boldsymbol{x} \int_0^{1/T} d\tau$. $\boldsymbol{x}$ denotes the position, τ is the imaginary-time variable, V and T are the system volume and the temperature, respectively, and $\sigma = \uparrow,\downarrow \equiv +,-$ is the spin label. We use units such that $\hbar = k_{\rm B} = 1$. $S_{\rm int}$ describes an electron-electron interaction via a two-body potential $v(\boldsymbol{x})$,

$$\begin{aligned} S_{\rm int} \ &= \ -\frac{1}{2}\int dx_1\, dx_2 \sum_{\sigma_1,\sigma_2} v(\boldsymbol{x}_1 - \boldsymbol{x}_2)\, \delta(\tau_1 - \tau_2) \\ &\quad \times \bar{\psi}_{\sigma_1}(x_1)\, \bar{\psi}_{\sigma_2}(x_2)\, \psi_{\sigma_2}(x_2)\, \psi_{\sigma_1}(x_1)\ . \end{aligned} \tag{2c}$$

In a microscopic theory the potential $v(\boldsymbol{x})$ would be the Coulomb interaction. Here we assume for simplicity that the theory has already been renormalized to take into account screening, so $v(\boldsymbol{x})$ represents a statically screened Coulomb interaction, or a some similar short-ranged model interaction. Equation (2c) describes the interactions of number-density fluctuations at all wavelengths, and in a microscopic description this is the only interaction term there is.[32] Integrating out fluctuations at short wavelengths in order to generate an effective long-wavelength theory generates interaction amplitudes in the spin-triplet channel in addition to the spin-singlet one, and in the particle-particle channel in addition to the particle-hole hole one, and all of these effective interaction amplitudes appear in all angular momentum channels.[16,33,34] The theory we will develop is general, and any desired interaction amplitudes can be kept. However, for the sake of transparency and simplicity of the formalism we will restrict ourselves to a finite number. Some of the most interesting applications of the theory are related to instabilities of the Fermi-liquid state due to a density-density interaction, and to magnetic order of the ferromagnetic or magnetic nematic variety. Furthermore, the soft modes in the particle-particle channel become massive in the presence of a magnetic field, as they do in disordered systems,[14] and hence can be suppressed experimentally. Accordingly, in the remainder of this paper we keep only interaction amplitudes in the particle-hole channel, and further restrict ourselves to the spin-singlet s-wave ($\ell = 0$) and the spin-triplet s-wave ($\ell = 0$) and p-wave ($\ell = 1$) channels. Generalizations to other interaction channels, for instance, the spin-triplet d-wave channel that has been discussed in the literature,[28] are straightforward. It needs to be stressed, however, that a restriction to low-angular-momentum channels represents an approximation over and above the low-energy restriction that is central to our approach, and that a complete low-energy effective theory needs to keep all angular momentum channels.

We define Fourier transforms of the fermionic field

$$\begin{aligned} \bar{\psi}_\sigma(k) \ &\equiv \ \bar{\psi}_{n\sigma}(\boldsymbol{k}) = \sqrt{T/V}\int dx\ e^{-ikx}\, \bar{\psi}_\sigma(x), \\ \psi_\sigma(k) \ &\equiv \ \psi_{n\sigma}(\boldsymbol{k}) = \sqrt{T/V}\int dx\ e^{ikx}\, \psi_\sigma(x), \end{aligned} \tag{3}$$

where $k = (\boldsymbol{k},\omega_n)$ is a four-vector that comprises a wave vector $\boldsymbol{k}$ and a fermionic Matsubara frequency $\omega_n = 2\pi T(n+1/2)$, and $kx \equiv \boldsymbol{k}\cdot\boldsymbol{x} - \omega_n\tau$. With the restrictions explained above the action then reads

$$S = \sum_{k,\sigma} \bar{\psi}_\sigma(k)\ \left[i\omega_n - \boldsymbol{k}^2/2m_{\rm e} + \mu\right]\ \psi_\sigma(k) + S_{\rm int}, \tag{4a}$$

with

$$\begin{aligned} S_{\rm int} &= -\frac{\Gamma_{\rm s}^{(0)}}{2}\,\frac{T}{V}{\sum_q}'\sum_{i=1}^{3} n(q)\,n(-q) \\ &\quad +\frac{\Gamma_{\rm t}^{(0)}}{2}\,\frac{T}{V}{\sum_q}' n_{\rm s}^{i}(q)\,n_{\rm s}^{i}(-q) \\ &\quad +\frac{\Gamma_{\rm t}^{(1)}}{2}\,\frac{T}{V}{\sum_q}'\sum_{i,\alpha=1}^{3} j_{\rm s}^{i\alpha}(q)\,j_{\rm s}^{i\alpha}(-q) \end{aligned} \tag{4b}$$

Here $q=(\boldsymbol{q},\Omega_n)$ comprises a wave vector $\boldsymbol{q}$ and a bosonic Matsubara frequency $\Omega_n=2\pi Tn$. $\Gamma_{\rm s}^{(0)}$, $\Gamma_{\rm t}^{(0)}$, and $\Gamma_{\rm s}^{(1)}$ are the interaction amplitudes in the particle-hole spin-singlet s-wave, spin-triplet s-wave, and spin-triplet p-wave channels, respectively. They are related to the parameters $F_0^{\rm s}$, $F_0^{\rm a}$, and $F_1^{\rm a}$, respectively, in Landau Fermi-liquid theory.[3] ${\sum_{\boldsymbol{q}}}'$ denotes a sum over wave vectors that is restricted to $|\boldsymbol{q}|<\Lambda$ with a cutoff wave number Λ. The long-wavelength properties we are interested in do not depend on Λ. The electron number density n, spin density $n_{\rm s}$, and spin current density $j_{\rm s}^{i\alpha}$ are given in Eqs. (5) below, and for completeness we also list the number current density j^{α}:

$$n(q) = \sum_k\sum_\sigma \bar\psi_\sigma(k+q/2)\,\psi_\sigma(k-q/2), \tag{5a}$$

$$j^{\alpha}(q) = \sum_k k_\alpha\sum_\sigma \bar\psi_\sigma(k+q/2)\,\psi_\sigma(k-q/2), \tag{5b}$$

$$n_{\rm s}^{i}(q) = \sum_k\sum_{\sigma_1,\sigma_2} \bar\psi_{\sigma_1}(k+q/2)\,(\sigma_i)_{\sigma_1\sigma_2}\,\psi_{\sigma_2}(k-q/2), \tag{5c}$$

$$j_{\rm s}^{i\alpha}(q) = \sum_k k_\alpha\sum_{\sigma_1,\sigma_2} \bar\psi_{\sigma_1}(k+q/2)(\sigma_i)_{\sigma_1\sigma_2}\psi_{\sigma_2}(k-q/2). \tag{5d}$$

The σ_i $(i=1,2,3)$ are the Pauli matrices, and the k_α $(\alpha=1,2,3)$ are the components of the wave vector $\boldsymbol{k}$. $\Gamma_{\rm s}^{(0)}>0$ describes a repulsive density-density interaction, and $\Gamma_{\rm t}^{0,1}>0$ describe ferromagnetic interactions in the s- and p-wave channels, respectively.

B. Bosonic formulation

1. *Phase-space degrees of freedom*

Our goal is to rewrite the action in terms of bosonic matrix fields. For this purpose it is convenient to define a bispinor[11]

$$\eta_n(\mathbf{x}) = \frac{1}{\sqrt{2}}\begin{pmatrix} \bar\psi_{n\uparrow}(\mathbf{x}) \\ \bar\psi_{n\downarrow}(\mathbf{x}) \\ \psi_{n\downarrow}(\mathbf{x}) \\ -\psi_{n\uparrow}(\mathbf{x}) \end{pmatrix} \quad , \tag{6a}$$

with an adjoint[35]

$$\begin{aligned} \eta_n^{+}(\boldsymbol{x}) &= (C\eta)_n(\boldsymbol{x}) \\ &= \frac{i}{\sqrt{2}}\left(-\psi_{n\uparrow}(\boldsymbol{x}),-\psi_{n\downarrow}(\boldsymbol{x}),\bar\psi_{n\downarrow}(\boldsymbol{x}),-\bar\psi_{n\uparrow}(\boldsymbol{x})\right). \end{aligned} \tag{6b}$$

Here $C_{nm}=i(\tau_1\otimes s_2)\,\delta_{nm}$ is the charge conjugation matrix in the spin-quaternion space spanned by $\tau_i\otimes s_j$ $(i,j=0,1,2,3)$ with $\tau_i=-s_j=-i\sigma_j$ with σ_0 the 2×2 unit matrix and $\sigma_{1,2,3}$ the Pauli matrices. Explicitly,

$$\tau_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad , \quad \tau_1 = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} \quad , \quad \tau_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} , \quad \tau_3 = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \quad , \tag{7a}$$

$$s_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad , \quad s_1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \quad , \quad s_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} , \quad s_3 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \quad . \tag{7b}$$

We now define a bilinear tensor product

$$\begin{aligned} B_{nm}(\boldsymbol{x},\boldsymbol{y}) &= \eta_n^+(\boldsymbol{x}) \otimes \eta_m(\boldsymbol{y}) \\ &= \frac{i}{2} \begin{pmatrix} -\psi_{n\uparrow}(\boldsymbol{x})\bar{\psi}_{m\uparrow}(\boldsymbol{y}) & -\psi_{n\uparrow}(\boldsymbol{x})\bar{\psi}_{m\downarrow}(\boldsymbol{y}) & -\psi_{n\uparrow}(\boldsymbol{x})\psi_{m\downarrow}(\boldsymbol{y}) & \psi_{n\uparrow}(\boldsymbol{x})\psi_{m\uparrow}(\boldsymbol{y}) \\ -\psi_{n\downarrow}(\boldsymbol{x})\bar{\psi}_{m\uparrow}(\boldsymbol{y}) & -\psi_{n\downarrow}(\boldsymbol{x})\bar{\psi}_{m\downarrow}(\boldsymbol{y}) & -\psi_{n\downarrow}(\boldsymbol{x})\psi_{m\downarrow}(\boldsymbol{y}) & \psi_{n\downarrow}(\boldsymbol{x})\psi_{m\uparrow}(\boldsymbol{y}) \\ \bar{\psi}_{n\downarrow}(\boldsymbol{x})\bar{\psi}_{m\uparrow}(\boldsymbol{y}) & \bar{\psi}_{n\downarrow}(\boldsymbol{x})\bar{\psi}_{m\downarrow}(\boldsymbol{y}) & \bar{\psi}_{n\downarrow}(\boldsymbol{x})\psi_{m\downarrow}(\boldsymbol{y}) & -\bar{\psi}_{n\downarrow}(\boldsymbol{x})\psi_{m\uparrow}(\boldsymbol{y}) \\ -\bar{\psi}_{n\uparrow}(\boldsymbol{x})\bar{\psi}_{m\uparrow}(\boldsymbol{y}) & -\bar{\psi}_{n\uparrow}(\boldsymbol{x})\bar{\psi}_{m\downarrow}(\boldsymbol{y}) & -\bar{\psi}_{n\uparrow}(\boldsymbol{x})\psi_{m\downarrow}(\boldsymbol{y}) & \bar{\psi}_{n\uparrow}(\boldsymbol{x})\psi_{m\uparrow}(\boldsymbol{y}) \end{pmatrix} \end{aligned} \tag{8}$$

and its Fourier transform

$$B_{nm}(\boldsymbol{k},\boldsymbol{p}) = \frac{1}{V} \int d\boldsymbol{x}\, d\boldsymbol{y}\; e^{-i\boldsymbol{k}\cdot\boldsymbol{x}+i\boldsymbol{p}\cdot\boldsymbol{y}}\, B_{nm}(\boldsymbol{x},\boldsymbol{y}). \tag{9a}$$

The 4×4 matrix $B_{nm}(\boldsymbol{k},\boldsymbol{p})$ can be expanded in the spin-quaternion basis defined above,

$$B_{nm}(\boldsymbol{k},\boldsymbol{p}) = \sum_{i,r=0}^{3} {}_r^i B_{nm}(\boldsymbol{k},\boldsymbol{p})\, (\tau_r \otimes s_i) \ . \tag{9b}$$

It is further useful to define

$$B_{nm}(\boldsymbol{k};\boldsymbol{q}) = B_{nm}(\boldsymbol{k}+\boldsymbol{q}/2, \boldsymbol{k}-\boldsymbol{q}/2) \ , \tag{9c}$$

with analogous definitions for other objects that depend on two wavevectors. All bilinear products of the fermion fields $\bar{\psi}$ and ψ can be written in terms of B, and in particular all terms in the interacting part of the action, Eq. (4b), can be written in terms of products of the B. An inspection shows that in the spin-quaternion basis, Eq. (7), the matrix elements ${}_{r=0,3}^{\ i} B$ and ${}_{r=1,2}^{\ i} B$ describe the particle-hole and particle-particle channels, respectively, and the ${}_r^{i=0} B$ describe the spin-singlet channel while the ${}_r^{i=1,2,3} B$ describe the spin-triplet channel. Taking moments of $B_{nm}(\boldsymbol{k};\boldsymbol{q})$ with respect to $\boldsymbol{k}$ generates variables in different angular momentum channels. For instance,

$$B_{nm}^{(0)}(\boldsymbol{q}) = \sum_{\boldsymbol{k}} B_{nm}(\boldsymbol{k};\boldsymbol{q}) \tag{10a}$$

defines s-wave or density degrees of freedom,

$$\boldsymbol{B}_{nm}^{(1)}(\boldsymbol{q}) = \sum_{\boldsymbol{k}} \boldsymbol{k}\, B_{nm}(\boldsymbol{k};\boldsymbol{q}) \tag{10b}$$

defines current degrees of freedom, etc.

We note that the introduction of bispinors, and of matrices with spin-quaternion valued matrix elements, is necessary in order to handle both particle-particle and particle-hole degrees of freedom within the same framework, as it allows to form bilinear products $\bar{\psi}\bar{\psi}$ and $\psi\psi$ in addition to $\bar{\psi}\psi$, see Eq. (8). For purposes that involve the particle-hole channel only one can restrict oneself to ordinary spinors, in which case the matrix B, Eq. (8), becomes a 2×2 matrix. Another observation is that B constitutes an overcomplete representation of bilinear fermion degrees of freedom since all of the matrix elements of B are not independent. We will come back to this point in Sec. II B 3 below.

2. Q-matrix field theory

Our next step is to constrain the matrices B in the interaction terms to a classical matrix field Q by means of a Lagrange multiplier field $\tilde{\Lambda}$. The fermion fields then enter the action only bilinearly and can be integrated out excatly.

This way we obtain an effective action $\mathcal{A}$ that depends on Q and $\tilde{\Lambda}$ according to

$$\begin{aligned} Z &= \int D[\bar{\psi},\psi]\; e^{S[\bar{\psi},\psi]} = \int D[\eta]\; e^{S[\eta]} \\ &= \int D[\eta]\; e^{S[\eta]} \int D[Q,\tilde{\Lambda}]\; e^{\mathrm{Tr}\,[\tilde{\Lambda}(Q-B)]} \\ &= \int D[Q,\tilde{\Lambda}]\; e^{\mathcal{A}[Q,\tilde{\Lambda}]}\ . \end{aligned} \tag{11}$$

Here and it what follows Tr denotes a trace over all degrees of freedom, including the continuous position in real space, while by tr we will denote a trace over all discrete degrees of freedom that are not explicitly shown.

In order to describe phenomena in angular-momentum channels higher than the s-wave one, or even in order to keep non-s-wave interaction constants, one obviously needs to apply this procedure to the phase-space variables defined in Eqs. (8, 9). As we will see later, this leads to a local field theory that allows for a systematic loop expansion about a saddle-point solution that describes a Fermi liquid. For certain purposes that require a density-channel interaction only it is technically advantageous to formulate the theory in terms of density variables, Eq. (10a), even though this leads to a nonlocal theory, i.e., one where the vertices diverge in the limit of small wave numbers. The starting point for such a formulation was given in Ref. 16, and we will revisit and further develop this method in Sec. IV E.

Application of the procedure sketched in Eq. (11) to the phase space variables, Eqs. (8, 9) yields the following formal expression for the effective action:

$$\mathcal{A}[Q,\tilde{\Lambda}] = \mathcal{A}_0 + \mathrm{Tr}\,\left(\tilde{\Lambda}^T Q\right) + \mathcal{A}_{\mathrm{int}}[Q]\ . \tag{12a}$$

Here

$$\mathcal{A}_0 = \frac{1}{2}\,\mathrm{Tr}\,\ln G^{-1}\ , \tag{12b}$$

where

$$G^{-1} = G_0^{-1} - i\tilde{\Lambda} \tag{12c}$$

is the inverse Green operator, with

$$\begin{aligned} \left(G_0^{-1}\right)_{nm}(\boldsymbol{x},\boldsymbol{y}) &= \delta_{nm}\;(\tau_0\otimes s_0)\;\delta(\boldsymbol{x}-\boldsymbol{y}) \\ &\quad\times\left[i\omega_n + \boldsymbol{\nabla}^2/2m_{\mathrm{e}} + \mu\right] \end{aligned} \tag{12d}$$

the bare Green operator. If a nontrivial band structure is desired, $-\boldsymbol{\nabla}^2/2m_{\mathrm{e}}$ should be replaced by an appropriate energy function $\epsilon(\boldsymbol{\nabla})$, see Eqs. (2). Rewriting the interaction part of the action, Eq. (4b), in terms of the B, and constraining the latter to Q by means of the functional delta-constraint we find

$$\mathcal{A}_{\mathrm{int}} = \mathcal{A}_{\mathrm{int}}^{(0,\mathrm{s})} + \mathcal{A}_{\mathrm{int}}^{(0,\mathrm{t})} + \mathcal{A}_{\mathrm{int}}^{(1,\mathrm{t})} \tag{13a}$$

where

$$\mathcal{A}_{\mathrm{int}}^{(0,\mathrm{s})} = \frac{T\,\Gamma_{\mathrm{s}}^{(0)}}{2V}\sum_{r,s=0,3}(-)^r \sum_{\substack{n_1,n_2\\ n_3,n_4}}\delta_{n_1-n_2,n_4-n_3}\sum_{\boldsymbol{k},\boldsymbol{p}}{\sum_{\boldsymbol{q}}}'\mathrm{tr}\left((\tau_r\otimes s_0)\,Q_{n_1n_2}(\boldsymbol{k};\boldsymbol{q})\right)\left(\tau_r\otimes s_0)\,Q_{n_3n_4}(\boldsymbol{p};-\boldsymbol{q})\right) \tag{13b}$$

$$\mathcal{A}_{\mathrm{int}}^{(0,\mathrm{t})} = \frac{T\,\Gamma_{\mathrm{t}}^{(0)}}{2V}\sum_{r,s=0,3}(-)^r \sum_{i=1}^{3}\sum_{\substack{n_1,n_2\\ n_3,n_4}}\delta_{N_1-n_2,n_4-n_3}\sum_{\boldsymbol{k},\boldsymbol{p}}{\sum_{\boldsymbol{q}}}'\mathrm{tr}\left((\tau_r\otimes s_i)\,Q_{n_1n_2}(\boldsymbol{k};\boldsymbol{q})\right)\left(\tau_r\otimes s_i)\,Q_{n_3n_4}(\boldsymbol{p};-\boldsymbol{q})\right) \tag{13c}$$

$$\mathcal{A}_{\mathrm{int}}^{(1,\mathrm{t})} = \frac{T\,\Gamma_{\mathrm{t}}^{(1)}}{2V}\sum_{r,s=0,3}(-)^r \sum_{i=1}^{3}\sum_{\substack{n_1,n_2\\ n_3,n_4}}\delta_{n_1-n_2,n_4-n_3}\sum_{\boldsymbol{k},\boldsymbol{p}}\boldsymbol{k}\cdot\boldsymbol{p}\,{\sum_{\boldsymbol{q}}}'\mathrm{tr}\left((\tau_r\otimes s_i)\,Q_{n_1n_2}(\boldsymbol{k};\boldsymbol{q})\right)\left(\tau_r\otimes s_i)\,Q_{n_3n_4}(\boldsymbol{p};-\boldsymbol{q})\right) \tag{13d}$$

3. Symmetry properties, and representation of observables

We now derive some useful symmetry properties of the Q-matrices. B as defined in Eq. (8) is self-adjoint under the operation defined in Eq. (6b). Q inherits this property, so we have

$$Q^+ = C^T\, Q^T\, C = Q \ . \tag{14a}$$

In the spin-quaternion basis defined in Eq. (7) this implies

$$ {}^{i}_{r}Q_{nm}(\boldsymbol{x},\boldsymbol{y}) \ = \ \begin{pmatrix} + \\ + \\ + \\ - \end{pmatrix}_r \begin{pmatrix} + \\ - \\ - \\ - \end{pmatrix}_i {}^{i}_{r}Q_{mn}(\boldsymbol{y},\boldsymbol{x}) \ , \tag{14b}$$

$$ {}^{i}_{r}Q_{nm}(\boldsymbol{k},\boldsymbol{p}) \ = \ \begin{pmatrix} + \\ + \\ + \\ - \end{pmatrix}_r \begin{pmatrix} + \\ - \\ - \\ - \end{pmatrix}_i {}^{i}_{r}Q_{mn}(-\boldsymbol{p},-\boldsymbol{k}) \ , \tag{14c}$$

$$ {}^{i}_{r}Q_{nm}(\boldsymbol{k};\boldsymbol{q}) \ = \ \begin{pmatrix} + \\ + \\ + \\ - \end{pmatrix}_r \begin{pmatrix} + \\ - \\ - \\ - \end{pmatrix}_i {}^{i}_{r}Q_{mn}(-\boldsymbol{k};\boldsymbol{q}) \ . \tag{14d}$$

Here the symbols $\begin{pmatrix} + \\ + \\ + \\ - \end{pmatrix}_r$ etc. denote a factor of $+1$ for $r = 0$ and a factor of -1 for $r = 1, 2, 3$, and analogously for i. We have made use of these relations in order to write the interacting part of the action in the form of Eqs. (13). They imply that all of the Q-matrix elements are not independent. In a model with N Matsubara frequencies, only $N(N+1)/2$ matrix elements are independent. We will later choose these to be the ones with $n \geq m$.

In terms of the Q-matrices the observables listed in Eqs. (5) are given by

$$ n(\boldsymbol{q}, i\Omega_n) \ = \ \sum_m \sum_{r=0,3} (-)^{3r/2} \sum_{\boldsymbol{k}} {}^{0}_{r}Q_{m,m+n}(\boldsymbol{k};\boldsymbol{q}) \ , \tag{15a}$$

$$ j_\alpha(\boldsymbol{q}, i\Omega_n) \ = \ \sum_m \sum_{r=0,3} (-)^{3r/2} \sum_{\boldsymbol{k}} k_\alpha \, {}^{0}_{r}Q_{m,m+n}(\boldsymbol{k};\boldsymbol{q}) \ , \tag{15b}$$

$$ n_{\rm s}^{i}(\boldsymbol{q}, i\Omega_n) \ = \ \sum_m \sum_{r=0,3} (-)^{3r/2} \sum_{\boldsymbol{k}} {}^{0}_{r}Q_{m,m+n}(\boldsymbol{k};\boldsymbol{q}) \ , \tag{15c}$$

$$ j_{\rm s}^{i\alpha}(\boldsymbol{q}, i\Omega_n) \ = \ \sum_m \sum_{r=0,3} (-)^{3r/2} \sum_{\boldsymbol{k}} k_\alpha \, {}^{0}_{r}Q_{m,m+n}(\boldsymbol{k};\boldsymbol{q}) \ . \tag{15d}$$

These are just examples of spin-singlet and spin-triplet observables in the $\ell = 0$ and $\ell = 1$ channels, respectively. Clearly, and desired observables can be expressed in terms of the Q.

4. Correlation functions

Physical correlation functions can be written in terms of Q-correlation functions by keeping appropriate source terms in the action while performing the transformation to the Q-matrix variables. For instance, for the Green function $G_{n\sigma}(\boldsymbol{x}-\boldsymbol{y}) = \langle \psi_{n\sigma}(\boldsymbol{x})\, \bar{\psi}_{n\sigma}(\boldsymbol{y})\rangle$ we obtain

$$ G_{n\sigma}(\boldsymbol{x}-\boldsymbol{y}) = \frac{i}{2}\, {\rm tr}\left[\big((\tau_0 + i\tau_3) \otimes (s_0 \mp \sigma i s_3)\big) \langle Q_{nn}(\boldsymbol{x},\boldsymbol{y})\rangle \right] \ , \tag{16}$$

with $\sigma = \uparrow, \downarrow \equiv +1, -1$. This can also be read off of Eq. (8) directly by keeping in mind the isomorphism between B and Q. For the density of states N as a function of the distance ω from the Fermi surface this implies

$$N(\omega) = \frac{4}{\pi} \, \mathrm{Re} \, \langle {}^0_0 Q_{nn}(\boldsymbol{x}, \boldsymbol{x}) \rangle \Big|_{i\omega_n \to \omega + i0} \; , \tag{17}$$

where we have used the symmetry properties expressed in Eqs. (14). Similarly, the number density susceptibility χ, and the spin density susceptibility tensor χ_{s}^{ij}, can be written

$$\begin{aligned} \chi(\boldsymbol{q}, i\Omega_n) \; &= \; 16\, T \sum_{m_1, m_2} \sum_{r=0,3} \sum_{\boldsymbol{k}, \boldsymbol{p}} \big\langle {}^0_r (\Delta Q)_{m_1 - n, m_1}(\boldsymbol{k}; \boldsymbol{q}) \\ &\qquad \times {}^i_r (\Delta Q)_{m_2, m_2 + n}(\boldsymbol{p}; -\boldsymbol{q}) \big\rangle \; , \end{aligned} \tag{18a}$$

$$\begin{aligned} \chi_{\mathrm{s}}^{ij}(\boldsymbol{q}, i\Omega_n) \; &= \; 16\, T \sum_{m_1, m_2} \sum_{r=0,3} \sum_{\boldsymbol{k}, \boldsymbol{p}} \big\langle {}^i_r (\Delta Q)_{m_1 - n, m_1}(\boldsymbol{k}; \boldsymbol{q}) \\ &\qquad \times {}^j_r (\Delta Q)_{m_2, m_2 + n}(\boldsymbol{p}; -\boldsymbol{q}) \big\rangle \; , \end{aligned} \tag{18b}$$

Here $\Delta Q = Q - \langle Q \rangle$, and, in Eq. (18b), $i, j = 1, 2, 3$. Other correlation functions can be expressed analogously.

III. IDENTIFICATION OF SOFT MODES: BROKEN SYMMETRY, AND A WARD IDENTITY

We are interested in separating the degrees of freedom represented by the Q-matrices into soft and massive modes. To identify the soft modes we derive a Ward identity that relates two-point Q-correlation functions to other quantities. We first do so for noninteracting electrons and then discuss the effects of interactions.

A. A Ward identity for noninteracting electrons

Let us consider transformations of the bispinors defined in Eqs. (6),

$$\eta_n(\boldsymbol{x}) \to \int d\boldsymbol{y} \; \hat{T}^{(\pm)}_{nm}(\boldsymbol{x}, \boldsymbol{y}) \, \eta_m(\boldsymbol{y}) \; , \tag{19a}$$

where the operator $\hat{T}^{(\pm)}$ defines non-local infinitesimal rotations in frequency space,

$$\hat{T}^{(\pm)}_{nm}(\boldsymbol{x}, \boldsymbol{y}) \; = \; (\tau_0 \otimes s_0) \, t^{(\pm)}_{nm}(\boldsymbol{x}, \boldsymbol{y}) \; , \tag{19b}$$

$$\begin{aligned} t^{(\pm)}_{nm}(\boldsymbol{x}, \boldsymbol{y}) \; &= \; \delta_{nm} \, \delta(\boldsymbol{x} - \boldsymbol{y}) + [\delta_{nn_1} \delta_{mn_2} - \delta_{nn_2} \delta_{mn_1}] \\ &\qquad \times \varphi_{\pm}(\boldsymbol{x}, \boldsymbol{y}) + O(\varphi^2) \; . \end{aligned} \tag{19c}$$

Here the

$$\begin{aligned} \varphi_{\pm}(\boldsymbol{x}, \boldsymbol{y}) \; &= \; \frac{1}{2} \, [\phi(\boldsymbol{x}, \boldsymbol{y}) \pm \phi(\boldsymbol{y}, \boldsymbol{x})] \\ &= \; \pm \varphi_{\pm}(\boldsymbol{y}, \boldsymbol{x}) \end{aligned} \tag{19d}$$

are even and odd combinations, respectively, of a non-local rotation angle ϕ. The matrices $\hat{T}^{\pm}$ obey $(\hat{T}^{(\pm)})^T = (\hat{T}^{(\pm)})^{-1}$, and $(\hat{T}^{(\pm)})^T \, C \, \hat{T}^{(\pm)} = C$ with C the charge conjugation matrix defined below Eq. (6b). For fixed n_1 and n_2 they form an SO(2) subgroup of a much larger symplectic group ($\mathrm{Sp}(8N, \mathbb{C})$ for a system with $2N$ frequency labels; N positive ones, including zero, and N negative ones), see Ref. 16. The Q-matrices transform as

$$Q(\boldsymbol{x}, \boldsymbol{y}) \to (T^{(\pm)} \, Q \, (T^{(\pm)})^T \; , \tag{20}$$

Explicitly we find

$$Q_{nm}(\boldsymbol{x}, \boldsymbol{y}) \to Q_{nm}(\boldsymbol{x}, \boldsymbol{y}) + \delta Q_{nm}(\boldsymbol{x}, \boldsymbol{y}) \tag{21a}$$

with

$$\delta Q_{nm}(\boldsymbol{x},\boldsymbol{y}) = \int d\boldsymbol{z}\,\Big\{\varphi_\pm(\boldsymbol{y},\boldsymbol{z})\;[\delta_{mn_1}\,Q_{nn_2}(\boldsymbol{x},\boldsymbol{z}) \mp \delta_{mn_2}\,Q_{nn_1}(\boldsymbol{x},\boldsymbol{z})] + \varphi_\pm(\boldsymbol{x},\boldsymbol{z})\;[\delta_{nn_1}\,Q_{n_2m}(\boldsymbol{z},\boldsymbol{y}) \mp \delta_{nn_2}\,Q_{n_1m}(\boldsymbol{z},\boldsymbol{y})]\Big\}\,, \tag{21b}$$

and in particular

$$\delta Q_{n_1n_2}(\boldsymbol{x},\boldsymbol{y}) = \int d\boldsymbol{z}\,\Big[\varphi_\pm(\boldsymbol{x},\boldsymbol{z})\,Q_{n_2n_2}(\boldsymbol{z},\boldsymbol{y}) - Q_{n_1n_1}(\boldsymbol{x},\boldsymbol{z})\,\varphi_\pm(\boldsymbol{z},\boldsymbol{y})\Big] \tag{21c}$$

The Lagrange multiplier field $\tilde{\Lambda}$ transforms as Q does, on account of the bilinear coupling between the two. Of the three terms in the action, Eqs. (12), the second one is invariant under these transformations, but $\mathcal{A}_0$ and $\mathcal{A}_{\rm int}$ are not. Focusing on noninteracting electrons for the time being, we find $\mathcal{A}_0 \to \mathcal{A}_0 + \delta\mathcal{A}_0$ with

$$\begin{aligned}\delta\mathcal{A}_0 \;&=\; \sum_{\boldsymbol{k},\boldsymbol{q}} [i\Omega_{n_1-n_2} + \boldsymbol{k}\cdot\boldsymbol{q}/m_{\rm e}]\;\mathrm{tr}\,G_{n_2n_1}(\boldsymbol{k};\boldsymbol{q}) \\ &\quad\times\varphi_\pm(\boldsymbol{k}-\boldsymbol{q}/2,\boldsymbol{k}+\boldsymbol{q}/2)\;.\end{aligned} \tag{22}$$

Here G is the inverse of the Green operator defined in Eq. (12c).

A Ward identity can now be derived by standard techniques.[7] We introduce a matrix source field J for the Q, consider the generating functional

$$Z[J] = \int D[Q,\tilde{\Lambda}]\;e^{\mathcal{A}_0 + \mathrm{Tr}\,(\tilde{\Lambda}^T Q) + \mathrm{Tr}\,(JQ)}\;, \tag{23}$$

perform the infinitesimal rotation defined by $\varphi_\pm$, differentiate with respect to J, and put $J=0$. This way we obtain a Ward identity

$$\langle \delta\mathcal{A}_0\,Q_{n_1n_2}(\boldsymbol{x},\boldsymbol{y})\rangle_{\mathcal{A}_0} + \langle \delta Q_{n_1n_2}(\boldsymbol{x},\boldsymbol{y})\rangle_{\mathcal{A}_0} = 0\;. \tag{24}$$

From Eqs. (22) and (24) we see that this relates correlation functions of the structure $\langle \mathrm{tr}\,G\,Q\rangle$ to $\langle Q\rangle$. The former can be rewritten in terms of $\langle Q\,Q\rangle$ by generalizing the generating functional given in Eq. (23). Since the Q are isomorphic to B, Eq. (8), we can write the source term $JQ = xJQ + (1-x)JB$ with an arbitrary real number x. The generating functional then becomes

$$Z[J] = \int D[Q,\tilde{\Lambda}]\;e^{\frac{1}{2}\,\mathrm{Tr}\,\ln[G^{-1} + i(1-x)J^T] + x\,\mathrm{Tr}\,(JQ) + \mathrm{Tr}\,(\tilde{\Lambda}^T Q)}\;. \tag{25}$$

Note that this is independent of x, and that by choosing $x=1$ we recover Eq. (23). By differentiating with respect to J, choosing $x=0$ and $x=1$, respectively, and putting $J=0$ we obtain an identity

$$\langle G_{n_2n_1}(\boldsymbol{x}_2,\boldsymbol{x}_1)\rangle = -2i\,\langle Q_{n_1n_2}(\boldsymbol{x}_1,\boldsymbol{x}_2)\rangle\;. \tag{26a}$$

Differentiating twice with respect to J we find

$$\langle G_{n_2n_1}(\boldsymbol{x}_2,\boldsymbol{x}_1)\,Q_{n_3n_4}(\boldsymbol{x}_3,\boldsymbol{x}_4)\rangle = -2i\,\langle Q_{n_1n_2}(\boldsymbol{x}_1,\boldsymbol{x}_2)\,Q_{n_3n_4}(\boldsymbol{x}_3,\boldsymbol{x}_4)\rangle\;. \tag{26b}$$

We now differentiate Eq. (24) with respect to ϕ in Eq. (19d). This yields two identities, one for φ_+ and one for φ_-. Adding them yields the Ward identity in its final form:

$$D_{n_1n_2,n_3n_4}(\boldsymbol{k},\boldsymbol{p};\boldsymbol{q}) \equiv \big\langle {}^0_0Q_{n_1n_2}(\boldsymbol{k};\boldsymbol{q})\,{}^0_0Q_{n_3n_4}(\boldsymbol{p};\boldsymbol{q})\big\rangle = \frac{i}{8}\,\delta_{\boldsymbol{k},\boldsymbol{p}}\;\delta_{n_1n_3}\,\delta_{n_2n_4}\,\frac{\big\langle {}^0_0Q_{n_1n_1}(\boldsymbol{k}-\boldsymbol{q}/2)\big\rangle - \big\langle {}^0_0Q_{n_2n_2}(\boldsymbol{k}+\boldsymbol{q}/2)\big\rangle}{i\Omega_{n_1-n_2} + \boldsymbol{k}\cdot\boldsymbol{q}/m_{\rm e}}\;. \tag{27}$$

B. Soft modes in noninteracting electron systems

Let us discuss the Ward identity for noninteracting electrons, Eq. (27). By using Eq. (26a) we see that the right-hand side of Eq. (27) is determined by

$$-2i\,\langle Q_{nn}(\boldsymbol{k}\pm\boldsymbol{q}/2)\rangle = \langle G_{nn}(\boldsymbol{k}+\boldsymbol{q}/2)\rangle = \frac{1}{i\omega_n - \xi_{\boldsymbol{k}\pm\boldsymbol{q}/2}} \to \mp i\pi\,\delta(\xi_{\boldsymbol{k}\pm\boldsymbol{q}/2}) - 1/\xi_{\boldsymbol{k}\pm\boldsymbol{q}/2}\;. \tag{28}$$

Here $\xi_{\boldsymbol{k}} = \boldsymbol{k}^2/2m_{\rm e} - \mu$, and the last term on indicates the limiting value of the Green function as ω_n approaches zero from above or below, respectively. We see that if n_1 and n_2 have opposite signs, then the numerator on the right-hand side of Eq. (27) goes to zero as $\Omega_{1-2} \to 0$ and $\boldsymbol{q} \to 0$, as does the denominator, whereas the numerator remains finite if n_1 and n_2 have opposite signs. In the latter case, the leading term for small Ω_{1-2} and small $\boldsymbol{q}$ is

$$\begin{aligned} &\left\langle {}^0_0 Q_{n_1 n_2}(\boldsymbol{k};\boldsymbol{q})\, {}^0_0 Q_{n_1 n_2}(\boldsymbol{p};-\boldsymbol{q})\right\rangle \;= \\ &\quad \delta_{\boldsymbol{k},\boldsymbol{p}}\, \frac{1}{16}\, \frac{2\pi i\, \delta(\xi_{\boldsymbol{k}})\, \mathrm{sgn}\, \Omega_{n_1-n_2}}{i\Omega_{n_1-n_2} + \boldsymbol{k}\cdot\boldsymbol{q}/m_{\rm e}} + \; O(1)\;. \end{aligned} \tag{29}$$

We see that there is an infinite number of soft modes that can be obtained by taking all possible moments of Eq. (29) with respect to the center-of-mass wave vector $\boldsymbol{k}$. This identifies the matrix elements Q_{nm} with $nm < 0$ as soft degrees of freedom, whereas the Q_{nm} with $nm > 0$ are massive. Physically, the soft modes are particle-hole excitations with a linear frequency-momentum relation. In the absence of interactions they are massless at nonzero temperature as well as at $T = 0$.

Equation (29) is a generalization of the Ward identity derived in Ref. 16, which corresponded to the zeroth moment at $\boldsymbol{q} = 0$,

$$\left\langle {}^0_0 Q^{(0)}_{n_1 n_2}(\boldsymbol{q})\, {}^0_0 Q^{(0)}_{n_1 n_2}(-\boldsymbol{q})\right\rangle\Big|_{\boldsymbol{q}=0} = \frac{\pi N_{\rm F}}{8|\Omega_{n_1-n_2}|}\;. \tag{30}$$

Here $Q^{(0)}(\boldsymbol{q}) = \sum_{\boldsymbol{k}} Q(\boldsymbol{k};\boldsymbol{q})$ is defined in analogy to $B^{(0)}$, Eq. (10a), and $N_{\rm F}$ is the free-electron density of states per spin at the Fermi level. Equation (30) is the clean analog of an identity that has been discussed before for disordered electrons.[9,10,36–38] In the presence of quenched disorder this is the *only* soft mode; all higher moments acquire a mass, see Appendix A. This is the reason why it is much harder to construct a soft-mode theory for clean electrons than for disordered ones: In the former case there are many more soft modes.

The Ward identity shows only that the matrix elements ${}^0_0 Q_{nm}$ with $nm < 0$ are soft. However, it is easy to see that the same statements holds for the ${}^i_r Q$ with arbitrary values of i and r. The reason is that the $\langle {}^i_r Q\, {}^i_r Q\rangle$ correlation functions are related to $\langle {}^0_0 Q\, {}^0_0 Q\rangle$ by means of symmetries that are not broken. This argument is identical to the one give in Ref. 16 and we do not repeat it here.

C. Interaction effects

We now need to address the question of what happens if the electron-electron interactions are taken into account. To this end we need to add the variation of the interacting part of the action, $\delta\mathcal{A}_{\rm int}$, under the transformation given in Eqs. (19), to Eq. (24). $\delta\mathcal{A}_{\rm int}$ is quadratic in Q, and the Ward identity now relates the 2-point function on the left-hand side of Eq. (27) to a 1-point function and a 3-point function. The technical details are cumbersome and have been derived in Ref. 16, but to show that the particle-hole excitations identified in Eq. (27) remain soft it suffices to discuss the structure of the generalized Ward identity, as was shown in Ref. 39. In what follows we recapitulate and adapt the argument give there. Equation (27) gets generalized to

$$\begin{aligned} &-8i\,(i\Omega_{n_1-n_2} + \boldsymbol{k}\cdot\boldsymbol{q}/m^*_{\rm e})\, D_{n_1 n_2, n_3 n_4}(\boldsymbol{k},\boldsymbol{p};\boldsymbol{q}) \\ &\quad = \delta_{n_1 n_3}\, \delta_{n_2 n_4}\, \delta_{\boldsymbol{k},\boldsymbol{p}}\, N_{n_1 n_2}(\boldsymbol{k},\boldsymbol{q}) - W_{n_1 n_2, n_3 n_4}(\boldsymbol{k},\boldsymbol{p};\boldsymbol{q})\;. \end{aligned} \tag{31a}$$

Here $m^*_{\rm e}$ is the effective electron mass that is renormalized by the Fermi-liquid parameter $F^1_{\rm s}$, and

$$N_{n_1 n_2}(\boldsymbol{k},\boldsymbol{q}) = \pi\, Z_{\boldsymbol{k}}\, \delta(\xi_{\boldsymbol{k}})\, \mathrm{sgn}\, \Omega_{n_1-n_2} \tag{31b}$$

with $Z_{\boldsymbol{k}}$ the quasi-particle weight that reflects the change from free-electron Green functions to physical ones. W is the contribution from $\delta\mathcal{A}_{\rm int}$. General considerations show that there are two contributions to D that are characterized by different frequency structures,

$$\begin{aligned} D_{n_1 n_2, n_3 n_4}(\boldsymbol{k},\boldsymbol{p};\boldsymbol{q}) \;=\; &\delta_{n_1 n_3}\, \delta_{n_2 n_4}\, D^{(\rm dc)}_{n_1 n_2}(\boldsymbol{k},\boldsymbol{p};\boldsymbol{q}) \\ &+\delta_{n_1-n_2, n_3-n_4}\, D^{(\rm c)}_{n_1 n_2, n_3 n_4}(\boldsymbol{k},\boldsymbol{p};\boldsymbol{q})\;, \end{aligned} \tag{32}$$

and analogously for W. Here the superscripts (dc) and (c), refer to the disconnected and connected contributions, respectively, in a representation in terms of fermionic diagrams. These contributions must respect the Ward identity

separately, and either one diverging in the limit of small frequencies and waves numbers will ensure the existence of soft modes. It therefore suffices to discuss $D^{(\mathrm{dc})}$. The Ward identity for the latter reads

$$\begin{aligned} -8i\,(i\Omega_{n_1-n_2} + \boldsymbol{k}\cdot\boldsymbol{q}/m_{\mathrm{e}}^*)\,D^{(\mathrm{dc})}_{n_1n_2}(\boldsymbol{k},\boldsymbol{p};\boldsymbol{q}) \;=\; \\ \delta_{\boldsymbol{k},\boldsymbol{p}}\,N_{n_1n_2}(\boldsymbol{k},\boldsymbol{q}) - W^{(\mathrm{dc})}_{n_1n_2}(\boldsymbol{k},\boldsymbol{p};\boldsymbol{q})\;. \end{aligned} \tag{33}$$

Now analytically continue to real frequencies $\Omega \in \mathbb{R}$, $i\Omega_{n_1-n_2} \to \Omega + i0$, and consider the limit $\Omega \to 0$, $q \to 0$. The only way for $D^{(\mathrm{dc})}$ at $t = 0$ to remain finite in this limit is for N and $W^{(\mathrm{dc})}$ at $\Omega = q = 0$ to cancel. However this is in general not possible: N approaches a nonzero limit for $\Omega \to 0$, $q \to 0$, and vanishing interactions, whereas W vanishes for noninteracting electrons. These two objects are thus different functions of the interaction, and their values at zero frequency and wave number can vanish at most for special values of the interaction. Notice that no such a cancellation is necessary to give $D^{(\mathrm{dc})}$ a mass at nonzero temperature. For simplicity, consider the zeroth moment with respect to $\boldsymbol{k}$ and $\boldsymbol{p}$ only, and put $\boldsymbol{q} = 0$. Then we have

$$-8i\Omega D^{(\mathrm{dc})} = N - W^{(\mathrm{dc})}\;. \tag{34}$$

Now let

$$D^{(\mathrm{dc})} \propto \frac{1}{\Omega + i/\tau_\phi}\;, \tag{35}$$

with $1/\tau_\phi$ a phase relaxation rate. As long as $\Omega\tau_\phi \gg 1$ for $T \to 0$ and $\Omega \to 0$, the right-hand side of Eq. (34) can be expanded in powers of $1/\Omega\tau_\phi$, with the leading contribution being a constant. As long as $1/\tau_\phi \to 0$ for $T \to 0$ such a regime always exists.

We conclude that the soft modes identified in Sec. III A, namely, the Q_{nm} with $nm < 0$, remain soft in the presence of interactions. The Ward identity shows this for the matrix elements ${}^0_0Q$ in the spin-quaternion basis, but additional symmetries of the action ensure that it is true for all of the ${}^i_rQ$.[16] This is consistent with both perturbation theory and with Fermi-liquid theory, which ensures that the number and nature of the soft modes cannot change as long as the system remains in a Fermi-liquid phase. The symmetry is broken, and the Goldstone modes exist, as long as there is a non-vanishing quasi-particle weight $Z_{\boldsymbol{k}}$. The only way the soft-mode spectrum can change is for the quasi-particle weight to vanish, which is to say that the system enters a non-Fermi-liquid state. We note that these considerations do not preclude using the soft-mode theory to study a breakdown of the Fermi liquid, which is signalized by a vanishing quasi-particle weight. They just indicate that a non-Fermi-liquid phase, where the symmetry is restored, has different properties and requires a different effective theory than the Fermi-liquid phase, but the theory we are about to construct will be valid everywhere in the latter, up to any possible transition where the quasi-particle weight vanishes. This is important for certain applications of the theory.[30]

We finally note that while the right-hand side of Eq. (29) resembles the structure of the Lindhard function, the soft modes are not related to particle-number conservation, and are not restricted to the density channel. Indeed, the density susceptibility cannot even be constructed from the left-hand side, as doing so would require *three* independent frequencies, rather than two, as can be seen from Eq. (18a). Rather, the infinitely many soft modes are the Goldstone modes of a spontaneously broken continuous symmetry, namely, the rotations in frequency space expressed by Eqs. (19).

IV. EFFECTIVE SOFT-MODE THEORY

A. Saddle-point solutions

We now return to the action, Eqs. (12, 13), and consider saddle-point solutions. There are many saddle-point solutions that have different symmetry properties and describe different physical states. Here we give just two examples. The first one is a saddle point that describes a Fermi liquid, and the second one reproduces the Stoner theory of ferromagnetism. Saddle points that describe superconducting states, magnetic states with non-s-wave order parameters, or, if a nontrivial band structure is taken into account, antiferromagnetic states, can be constructed analogously.

1. Fermi-liquid saddle point

The saddle-point equations are obtained by minimizing the action with respect to $\tilde{\Lambda}$ and Q. They read

$$0 = -\frac{i}{2}\left(G^{\rm sp}_{mn}(\boldsymbol{y},\boldsymbol{x})\right)^T + Q^{\rm sp}_{nm}(\boldsymbol{x},\boldsymbol{y}) \ . \tag{36a}$$

$$0 = \tilde{\Lambda}^{\rm sp}_{nm}(\boldsymbol{x},\boldsymbol{y}) + \frac{\delta}{\delta Q_{nm}(\boldsymbol{x},\boldsymbol{y})}\Big|_{Q^{\rm sp}} \mathcal{A}_{\rm int}[Q] \ . \tag{36b}$$

To find a saddle-point solution that describes a Fermi liquid we make an ansatz

$$\begin{aligned} Q^{\rm sp}_{nm}(\boldsymbol{x},\boldsymbol{y}) &= (\tau_0 \otimes s_0)\,\delta_{nm}\,Q_n(\boldsymbol{x}-\boldsymbol{y}) \ , \\ \tilde{\Lambda}^{\rm sp}_{nm}(\boldsymbol{x},\boldsymbol{y}) &= (\tau_0 \otimes s_0)\,\delta_{nm}\,\Lambda_n(\boldsymbol{x}-\boldsymbol{y}) \ . \end{aligned} \tag{37}$$

Performing the derivative in Eq. (36a) and going into Fourier space we find Q_n and Λ_n as the solution of the equations

$$Q_n(\boldsymbol{k}) = \frac{i}{2}\,\frac{1}{i\omega_n - \xi_{\boldsymbol{k}} - i\Lambda} \ , \tag{38a}$$

$$\Lambda_n(\boldsymbol{k}) \equiv \Lambda = -4\Gamma^{(0)}_{\rm s}\,\frac{1}{V}\sum_{\boldsymbol{p}} T\sum_m e^{i\omega_m 0}\,Q_m(\boldsymbol{p}) \ . \tag{38b}$$

We see that $-2iQ_n(\boldsymbol{k})$ equals the saddle-point Green function, in agreement with Eq. (26a), with Λ the self energy. The latter represents the spin-singlet interaction $\Gamma^{(0)}_{\rm s}$ in Hartree-Fock approximation. The factor $e^{i\omega_m 0}$ in Eq. (38b) is the usual convergence factor that resolves the ambiguity inherent in the equal-time Green function.[4] This Hartree-Fock saddle-point solution is a generalization of the saddle-point solution considered in Ref. 16 to a non-local Green function.

2. Stoner saddle point

To illustrate the existence of other saddle points, and the general versatility of the theory, we now consider a saddle-point ansatz

$$\begin{aligned} Q^{\rm sp}_{nm}(\boldsymbol{x},\boldsymbol{y}) &= \delta_{12}\Big[(\tau_0 \otimes s_0)\,G_n(\boldsymbol{x}-\boldsymbol{y}) \\ &\qquad +(\tau_3 \otimes s_3)\,F_n(\boldsymbol{x}-\boldsymbol{y})\Big] \ , \end{aligned} \tag{39a}$$

$$\begin{aligned} \tilde{\Lambda}^{\rm sp}_{nm}(\boldsymbol{x},\boldsymbol{y}) &= \delta_{nm}\Big[-(\tau_0 \otimes s_0)\,i\Sigma_n(\boldsymbol{x}-\boldsymbol{y}) \\ &\qquad +(\tau_3 \otimes s_3)\,i\Delta_n(\boldsymbol{x}-\boldsymbol{y})\Big] \ . \end{aligned} \tag{39b}$$

The saddle-point equations (36) now yield

$$G_n(\boldsymbol{k}) = \frac{i}{4}\left[\mathcal{G}^+_n(\boldsymbol{k}) + \mathcal{G}^-_n(\boldsymbol{k})\right] \ , \tag{40a}$$

$$F_n(\boldsymbol{k}) = \frac{i}{4}\left[\mathcal{G}^+_n(\boldsymbol{k}) - \mathcal{G}^-_n(\boldsymbol{k})\right] \ , \tag{40b}$$

$$\Sigma_n(\boldsymbol{k}) \equiv \Sigma = -4i\Gamma^{(0)}_{\rm s}\,\frac{1}{V}\sum_{\boldsymbol{p}} T\sum_m e^{i\omega_m 0}\,G_m(\boldsymbol{p}) \ , \tag{40c}$$

$$\Delta_n(\boldsymbol{k}) \equiv \Delta = -4i\Gamma^{(0)}_{\rm t}\,\frac{1}{V}\sum_{\boldsymbol{p}} T\sum_m e^{i\omega_m 0}\,F_m(\boldsymbol{p}) \ . \tag{40d}$$

Here

$$\mathcal{G}_n^{\pm}(\boldsymbol{k}) = \frac{1}{i\omega_n - \xi_{\boldsymbol{k}} \pm \Delta - \Sigma} \tag{41}$$

are Green functions whose self energy contributions describe the particle-hole spin-triplet and spin-singlet interactions in Hartree-Fock approximation. This saddle-point solution is the clean limit of the on considered in Ref. 40, generalized to the case of non-local Green functions. It is equivalent to Stoner theory, which can be seen as follows. $G_n(\boldsymbol{k})$ and $F_n)\boldsymbol{k})$ obey the equations

$$(i\omega_n - \xi_{\boldsymbol{k}} - \Sigma)\,\mathcal{G}_n(\boldsymbol{k}) + \Delta\,\mathcal{F}_n(\boldsymbol{k}) = 1\ , \tag{42a}$$

$$(i\omega_n - \xi_{\boldsymbol{k}} - \Sigma)\,\mathcal{F}_n(\boldsymbol{k}) + \Delta\,\mathcal{G}_n(\boldsymbol{k}) = 0\ . \tag{42b}$$

If we absorb the constant spin-singlet self energy Σ into a redefinition of the chemical potential we obtain from Eqs. (42) the equation of state in a form that is familiar from Stoner theory,[41]

$$1 = -2\Gamma_{\rm t}^{(0)}\,T\sum_n \frac{1}{V}\sum_{\boldsymbol{k}} \frac{1}{(i\omega_n - \xi_{\boldsymbol{k}})^2 - \Delta^2}\ , \tag{43}$$

which allows for a nonzero magnetization (which is proportional to Δ) provided $N_{\rm F}\Gamma_{\rm t}^{(0)} > 1$.

B. Expansion about the saddle point

In the remainder of this paper we consider fluctuations about the Fermi-liquid saddle point derived in Sec. IV A 1; an application of the theory to magnetically ordered states will be discussed elsewhere.[42]

1. Gaussian fluctuations

We first consider Gaussian fluctuations about the Fermi-liquid saddle point. If we write $Q = Q^{\rm sp} + \delta Q$ and $\tilde{\Lambda} = \tilde{\Lambda}^{\rm sp} + \delta\tilde{\Lambda}$, an expansion of the action, Eqs. (12, 13), yields

$$\mathcal{A}[Q,\tilde{\Lambda}] = \mathcal{A}^{\rm sp} + \mathcal{A}^{(2)} + \Delta\mathcal{A}\ , \tag{44}$$

where $\mathcal{A}^{\rm sp}$ is the saddle-point action, $\mathcal{A}^{(2)}$ denotes the Gaussian fluctuations, and $\Delta\mathcal{A}$ denotes the contributions that are of cubic or higher order in the fluctuations. For the Gaussian part we find

$$\mathcal{A}^{(2)} = \frac{1}{4}\,{\rm Tr}\,(G_{\rm sp}\,\delta\tilde{\Lambda}\,G_{\rm sp}\,\delta\tilde{\Lambda}) + {\rm Tr}\,(\delta\tilde{\Lambda}^T\delta Q) + \mathcal{A}_{\rm int}[\delta Q]\ . \tag{45}$$

Here

$$(G_{\rm sp})_{nm}\,(\boldsymbol{k},\boldsymbol{p}) = (\tau_0\otimes s_0)\,\delta_{nm}\,\delta_{\boldsymbol{k},\boldsymbol{p}}\,G_n(\boldsymbol{k}) \tag{46a}$$

with

$$G_n(\boldsymbol{k}) = \frac{1}{i\omega_n - \xi_{\boldsymbol{k}} - i\Lambda}\ , \tag{46b}$$

where Λ is the solution of Eqs. (38), is the Green function in saddle-point approximation. With our local interaction amplitude Λ amounts only to a constant shift of the chemical potential, so $G_n(\boldsymbol{k})$ can be taken to be the free-electron Green function.

We can decouple $\delta\tilde{\Lambda}$ and δQ at the Gaussian level by defining a field $\delta\bar{\Lambda}$ by

$$\overset{i}{_r}(\delta\bar{\Lambda})_{12} = \frac{1}{2}\,\phi_{12}\ \overset{i}{_r}(\delta\tilde{\Lambda})_{12} + \overset{i}{_r}(\delta\bar{Q})_{12}\ . \tag{47a}$$

Here we have defined $(\delta\bar{\Lambda})_{12} \equiv (\delta\bar{\Lambda})_{n_1n_2}(\boldsymbol{p}_1,\boldsymbol{p}_2)$, and analogously for other fields, as well as

$$\overset{i}{_r}(\delta\bar{Q})_{12} = \overset{i}{_r}(\delta Q)_{12}\ \begin{pmatrix} + \\ + \\ - \\ - \end{pmatrix}_r \begin{pmatrix} + \\ - \\ + \\ - \end{pmatrix}_i \tag{47b}$$

and

$$\begin{aligned}\phi_{12} &= G_1\, G_2 \equiv G_{n_1}(\boldsymbol{p}_1)\, G_{n_2}(\boldsymbol{p}_2) \\ &\equiv \begin{cases}\Phi_{12} & \text{if } (n_1+1/2)(n_2+1/2)>0 \\ \varphi_{12} & \text{if } (n_1+1/2)(n_2+1/2)<0\ .\end{cases}\end{aligned} \tag{47c}$$

The frequency-restricted objects Φ_{12} and φ_{12} will be useful later. We also define

$$(\delta Q)^{\ddagger}_{12} = (\delta Q)_{n_1 n_2}(-\boldsymbol{p}_1, -\boldsymbol{p}_2)\ , \tag{47d}$$

and analogously for $\delta\Lambda$. In terms of $\delta\bar{\Lambda}$ and δQ the Gaussian action then takes the form

$$\begin{aligned}\mathcal{A}^{(2)} &= -4\sum_{r,i}\begin{pmatrix}+\\-\\-\\+\end{pmatrix}_r \sum_{12}{}^{i}_{r}(\delta Q)_{12}\,\frac{1}{\phi_{12}}\,{}^{i}_{r}(\delta Q)^{\ddagger}_{12} \\ &\quad +4\sum_{r,i}\begin{pmatrix}+\\-\\-\\+\end{pmatrix}_r \sum_{12}{}^{i}_{r}(\delta\bar{\Lambda})_{12}\,\frac{1}{\phi_{12}}\,{}^{i}_{r}(\delta\bar{\Lambda})^{\ddagger}_{12}\ . \\ &\quad +\mathcal{A}_{\rm int}[\delta Q]\ .\end{aligned} \tag{48}$$

Note that the Gaussian $\delta\bar{\Lambda}$ propagator equals minus the δQ propagator[43] for noninteracting electrons,

$$\langle \delta\bar{\Lambda}_{12}\,\delta\bar{\Lambda}_{34}\rangle_{\mathcal{A}^{(2)}} = -\langle \delta Q_{12}\,\delta Q_{34}\rangle_{\mathcal{A}_0^{(2)}}\ , \tag{49}$$

with $\mathcal{A}_0^{(2)}$ the first two terms in Eq. (45) or (48). This will be important for the structure of the effective soft-mode theory and the loop expansion.

2. *The fluctuation action in terms of independent variables*

To go beyond Gaussian fluctuations we expand the Tr ln term in the action, Eq. (12), in powers of $\delta\tilde{\Lambda}$ and express the result in terms of $\delta\bar{\Lambda}$ and $\delta\bar{Q}$ by means of Eq. (47a). We find

$$\begin{aligned}\Delta\mathcal{A} &= -\sum_{m=3}^{\infty}\frac{i^m}{2m}\,{\rm Tr}\,(G_{\rm sp}\,\delta\tilde{\Lambda})^m \equiv \sum_{m=3}^{\infty}\Delta\mathcal{A}^{(m)} \\ &= \frac{4i}{3}\sum_{1,2,3} G_1^{-1}\,G_2^{-1}\,G_3^{-1}\,{\rm tr}\,\left[(\delta\bar{\Lambda}_{12}-\delta\bar{Q}_{12})(\delta\bar{\Lambda}_{23}-\delta\bar{Q}_{23})(\delta\bar{\Lambda}_{31}-\delta\bar{Q}_{31})\right] \\ &\quad -2\sum_{1,2,3,4} G_1^{-1}\,G_2^{-1}\,G_3^{-1}\,G_4^{-1}\,{\rm tr}\,\left[(\delta\bar{\Lambda}_{12}-\delta\bar{Q}_{12})(\delta\bar{\Lambda}_{23}-\delta\bar{Q}_{23})(\delta\bar{\Lambda}_{34}-\delta\bar{Q}_{34})(\delta\bar{\Lambda}_{41}-\delta\bar{Q}_{41})\right]+\ldots\end{aligned} \tag{50}$$

Notice that the vertices are given in terms of products of inverse Green functions, which are well behaved in the limit of small frequencies of wave numbers. This is thus a local field theory, in contrast to the formulation in terms of density modes in Ref. 16.

We now recall two features of the theory. First, all of the matrix elements of Q and Λ are not independent. We can choose the Q_{nm} with $n \geq m$ as the independent ones, and express those with $n < m$ in terms of them by means of Eqs. (14). Second, we recall from Sec. III that the Q_{nm} are massive degrees of freedom if n and m have the same sign, and soft ones if n and m have opposite signs. We express these facts in our notation by choosing as independent matrix elements

$$(\delta\bar{Q})_{12} = \begin{cases}\bar{P}_{12} & \text{if } n_1 \geq n_2 \wedge (n_1+1/2)(n_2+1/2)>0 \\ q_{12} & \text{if } n_1 \geq 0 \wedge n_2 < 0\ ,\end{cases} \tag{51a}$$

$$(\delta\bar{\Lambda})_{12} = \begin{cases}\Lambda_{12} & \text{if } n_1 \geq n_2 \wedge (n_1+1/2)(n_2+1/2)>0) \\ \lambda_{12} & \text{if } n_1 \geq 0 \wedge n_2 < 0\ .\end{cases} \tag{51b}$$

In what follows we will absorb the frequency restrictions into the fields, so writing, e.g., q_{12} implies that $n_1 \geq 0$ and $n_2 < 0$, etc. For the Gaussian action we obtain

$$\begin{aligned}\mathcal{A}^{(2)} &= -8\sum_{r,i}\begin{pmatrix}+\\-\\-\\+\end{pmatrix}_r \sum_{12}\frac{1}{\phi_{12}}\left[{}^{i}_{r}q_{12}\,{}^{i}_{r}q^{\ddagger}_{12} + I_{12}\,{}^{i}_{r}\bar{P}_{12}\,{}^{i}_{r}\bar{P}^{\ddagger}_{12} - {}^{i}_{r}\lambda_{12}\,{}^{i}_{r}\lambda^{\ddagger}_{12} - I_{12}\,{}^{i}_{r}\Lambda_{12}\,{}^{i}_{r}\Lambda^{\ddagger}_{12}\right] \\ &\quad + \frac{16T\Gamma_{\rm s}^{(0)}}{V}\sum_{r=0,3}\sum_{\substack{1,2\\3,4}}\delta_{1-2,3-4}\left[{}^{0}_{r}q_{12}\,{}^{0}_{r}q^{\ddagger}_{34} + I_{12}\,{}^{0}_{r}\bar{P}_{12}\,{}^{0}_{r}\bar{P}^{\ddagger}_{34} + {}^{0}_{r}q_{12}\,{}^{0}_{r}\bar{P}^{\ddagger}_{34} + {}^{0}_{r}\bar{P}_{12}\,{}^{0}_{r}q^{\ddagger}_{34}\right] , \end{aligned} \tag{52}$$

where $I_{12} = 1 - \delta_{n_1 n_2}/2$. Here we keep only the particle-hole spin-singlet interaction; the other interaction channels provide analogous contributions to the Gaussian action. Notice that the massless and massive degrees of freedom are decoupled in the noninteracting part of the action, but are coupled by the interaction via terms that are bilinear in q and $\bar{P}$.

The contributions to $\Delta\mathcal{A}$ can also be expressed in terms of $\bar{P}$, q, Λ, and λ. Simple combinatorics yield for the cubic term

$$\begin{aligned}\Delta\mathcal{A}^{(3)} &= -4i\Big[\mathrm{Tr}\,\left[G^{-1}\left(\bar{P}-\Lambda\right)G^{-1}\left(\bar{P}+\bar{P}^{+}-\Lambda-\Lambda^{+}\right)G^{-1}\left(\bar{P}-\Lambda\right)^{+}\right] \\ &\quad + \mathrm{Tr}\,\left[G^{-1}\left(\bar{P}+\bar{P}^{+}-\Lambda-\Lambda^{+}\right)G^{-1}\left(\not{q}\,G^{-1}\not{q}^{+}+\not{q}^{+}G^{-1}\not{q}\right)\right]\Big] , \end{aligned} \tag{53}$$

or, schematically,

$$\Delta\mathcal{A}^{(3)} \propto G^{-3}\left[\left(\bar{P}-\Lambda\right)^3 + \left(\bar{P}-\Lambda\right)\not{q}^2\right] , \tag{4.18$'$}$$

and for the quartic one

$$\begin{aligned}\Delta\mathcal{A}^{(4)} &= -2\,\mathrm{Tr}\,\Big[4\,G^{-1}\left(\bar{P}-\Lambda\right)G^{-1}\left(\bar{P}-\Lambda\right)G^{-1}\left(\bar{P}-\Lambda\right)G^{-1}\left(\bar{P}-\Lambda\right)^{+} \\ &\qquad +4\,G^{-1}\left(\bar{P}-\Lambda\right)^{+}G^{-1}\left(\bar{P}-\Lambda\right)^{+}G^{-1}\left(\bar{P}-\Lambda\right)^{+}G^{-1}\left(\bar{P}-\Lambda\right) \\ &\qquad +3\,G^{-1}\left(\bar{P}-\Lambda\right)G^{-1}\left(\bar{P}-\Lambda\right)G^{-1}\left(\bar{P}-\Lambda\right)^{+}G^{-1}\left(\bar{P}-\Lambda\right)^{+} \\ &\qquad +3\,G^{-1}\left(\bar{P}-\Lambda\right)G^{-1}\left(\bar{P}-\Lambda\right)^{+}G^{-1}\left(\bar{P}-\Lambda\right)G^{-1}\left(\bar{P}-\Lambda\right)^{+}\Big] \\ &\quad -8\,\mathrm{Tr}\,\Big[G^{-1}(\bar{P}+\bar{P}^{+}-\Lambda-\Lambda^{+})\,G^{-1}\left(\bar{P}+\bar{P}^{+}-\Lambda-\Lambda^{+}\right)G^{-1}\,\left(\not{q}\,G^{-1}\not{q}^{+}+\not{q}^{+}G^{-1}\not{q}\right) \\ &\qquad +G^{-1}\left(\bar{P}+\bar{P}^{+}-\Lambda-\Lambda^{+}\right)G^{-1}\not{q}\,G^{-1}\left(\bar{P}+\bar{P}^{+}-\Lambda-\Lambda^{+}\right)G^{-1}\not{q}^{+}\Big] \\ &\quad -4\,\mathrm{Tr}\,\left[G^{-1}\not{q}\,G^{-1}\not{q}^{+}G^{-1}\not{q}\,G^{-1}\not{q}^{+}\right] , \end{aligned} \tag{54}$$

or, schematically,

$$\Delta\mathcal{A}^{(4)} \propto G^{-4}\,\left[\left(\bar{P}-\Lambda\right)^4 + \left(\bar{P}-\Lambda\right)^2\not{q}^2 + \not{q}^4\right] . \tag{4.19$'$}$$

Here we have defined

$$\not{q} = q - \lambda , \tag{55}$$

and $G^{-1}\bar{P}$ etc. denote matrix products with the matrix elements of G given by $G_{12} = \delta_{12}\,G_1\,(\tau_0 \otimes s_0)$. The matrix elements of $\bar{P}$ are ${}^{i}_{r}\bar{P}_{n_1 n_2}(\boldsymbol{p}_1, \boldsymbol{p}_2)$, see the left-hand side of Eq. (14c), and those of $\bar{P}^{+}$ are $\begin{pmatrix}+\\+\\+\\-\end{pmatrix}_r \begin{pmatrix}+\\-\\-\\-\end{pmatrix}_i {}^{i}_{r}\bar{P}_{n_2 n_1}(-\boldsymbol{p}_2, -\boldsymbol{p}_1)$, see the right-hand side of Eq. (14c). The matrix elements of Λ, Λ^{+}, q, q^{+}, and λ, λ^{+}, respectively, are given by analogous expressions. This implies for the expansion of the various matrices in the spin-quaternion basis

$$\bar{P}_{12} = \sum_{r=0}^{3}\sum_{i=0}^{3}{}^{i}_{r}\bar{P}_{12}\,(\tau_r \otimes s_i) , \tag{56a}$$

$$\bar{P}^{+}_{12} = \sum_{r=0}^{3}\sum_{i=0}^{3}{}^{i}_{r}\bar{P}^{\ddagger}_{21}\,(\tau^{+}_r \otimes s^{+}_i) , \tag{56b}$$

where

$$ {}^{i}_{r}\bar{P}^{\ddagger}_{12} = {}^{i}_{r}\bar{P}_{n_1 n_2}(-\boldsymbol{p}_1, -\boldsymbol{p}_2) \quad , \quad (r = 0, 3) \tag{56c} $$

in the particle-hole channel, and

$$ {}^{i}_{r}\bar{P}^{\ddagger}_{12} = -{}^{i}_{r}\bar{P}_{n_1 n_2}(-\boldsymbol{p}_1, -\boldsymbol{p}_2) \quad , \quad (r = 1, 2) \tag{56d} $$

in the particle-particle channel. τ_r^+ and s_i^+ are the hermitian conjugates of τ_r and s_i, respectively. Analogous relations hold for Λ, q, and λ and their adjoints. In all cases the frequency ordering restrictions explained after Eqs. (51) apply.

Expressions for the higher-order terms in $\Delta\mathcal{A}$ in terms of P, q, Λ, and λ can be constructed analogously. For instance, $\Delta\mathcal{A}^{(5)}$ contains terms that have the schematic structures

$$ \Delta\mathcal{A}^{(5)} \propto G^{-5}\left[(\bar{P}-\Lambda)^5 + (\bar{P}-\Lambda)^3 \not{q}^2 + (\bar{P}-\Lambda)\not{q}^4\right] . \tag{57} $$

C. Integrating out the massive modes

We now recall that $\bar{P}$ and Λ represent massive fluctuations, whereas q and λ represent soft ones, and our goal is to construct an effective theory in terms of the latter. We cannot simply neglect $\bar{P}$ and Λ since they couple to q and λ; rather, we need to integrate them out in an approximation that respects the Ward identity discussed in Sec. III. We let ourselves be guided in this process by an analysis of the $O(2)$ ϕ^4-theory in Appendix B. The procedure used below for our matrix theory is analogous to the one explained for the $O(2)$ model in the context of Eqs. (B11). To avoid unnecessary notational complexity, from here on we will restrict ourselves to a model with only one interaction amplitude, $\Gamma_{\rm s}^{(0)} \equiv \Gamma$, in the particle-hole spin-singlet channel. The spin-triplet ($i = 1, 2, 3$) and particle-particle ($r = 1, 2$) degrees of freedom then just represent noninteracting electrons, and in the remainder of this paper we will neglect them. That is, we keep only matrix elements with $r = 0, 3$ and $i = 0$. It will be obvious how to generalize the theory to the presence of interactions in other channels.[44]

1. *Screening of the interaction*

It is advantageous to first eliminate the coupling between $\bar{P}$ and q at the Gaussian level. We see from Eq. (52) that this can be achieved by a simple shift of $\bar{P}$. We define

$$ {}^{0}_{r}P_{12} = {}^{0}_{r}\bar{P}_{12} - \frac{2T\Gamma}{V}\sum_{3,4,5,6} M^{-1}_{12,34}\,\delta_{3-4,5-6}\,{}^{0}_{r}q_{56} \ . \tag{58a} $$

Here

$$ \begin{aligned} M^{-1}_{12,34} \ &= \ \delta_{13}\delta_{24}\Phi_{12} \\ &\quad + \delta_{1-2,3-4}\,\frac{2(T/V)\Gamma\,\Phi_{12}\Phi_{34}}{1 - 2T\Gamma\frac{1}{V}\sum_{5,6}\delta_{1-2,5-6}\Phi_{56}} \end{aligned} \tag{58b} $$

is the inverse of

$$ M_{12,34} = \delta_{13}\delta_{24}\Phi^{-1}_{12} - \frac{2T\Gamma}{V}\,\delta_{1-2,3-4} \ . \tag{58c} $$

Due to the frequency restrictions inherent in Φ_{12}, see Eq. (47c), the frequency sum in the denominator of Eq. (58b) has no hydrodynamic content and can be replaced by minus the static electron density susceptibility in Hartree-Fock approximation, $\chi_{\rm st}$ which in turn we can replace by $N_{\rm F}$:

$$ -\frac{T}{V}\sum_{5,6}\delta_{1-2,5-6}\,\Phi_{56} \approx \chi_{\rm st} \approx N_{\rm F} \ . \tag{58d} $$

Equation (58a) thus simplifies to

$$ {}^{0}_{r}P_{12} \;=\; {}^{0}_{r}\bar{P}_{12} - 2T\gamma\,\Phi_{12}\,\frac{1}{V}\sum_{3,4}\delta_{1-2,3-4}\,{}^{0}_{r}q_{34} \tag{59a} $$

$$ \equiv\; {}^{0}_{r}\bar{P}_{12} - {}^{0}_{r}(\hat{\Phi}\hat{\gamma}q)_{12} \tag{59b} $$

with

$$ \gamma = \Gamma/(1+2N_{\rm F}\Gamma) \tag{59c} $$

the statically screened interaction amplitude and operators

$$ \hat{\gamma}_{12,34} \;=\; (\tau_0\otimes s_0)\,2(T/V)\gamma\,\delta_{1-2,3-4}\;, \tag{59d} $$

$$ \hat{\Phi}_{12,34} \;=\; (\tau_0\otimes s_0)\,\delta_{13}\,\delta_{24}\,\Phi_{12}\;. \tag{59e} $$

Analogously we define $\hat{\Gamma}$ and $\hat{\varphi}$. These operators are all self-adjoint with respect to the matrix adjoint, i.e., $(\hat{\gamma}q)^{+} = \hat{\gamma}q^{+}$ etc.

In addition to decoupling $\bar{P}$ and q in the Gaussian action, this procedure introduces a new term quadratic in q that combines with the q^2 term in Eq. (52) to change the interaction amplitude Γ to γ. That is, the bilinear coupling between the massive modes and the soft ones leads to the screening of the interaction. The Gaussian action in terms of P, Λ, q, and λ now reads

$$ \begin{aligned} \mathcal{A}^{(2)} \;=\; & -8\sum_{r=0,3}\sum_{\substack{1,2\\3,4}}\left[{}^{0}_{r}q_{12}\left(\delta_{13}\delta_{24}\,\frac{1}{\varphi_{12}} - \delta_{1-2,3-4}\,\frac{2T\gamma}{V}\right){}^{0}_{r}q^{\ddagger}_{34} - \delta_{13}\delta_{24}\,\frac{1}{\varphi_{12}}\,{}^{0}_{r}\lambda_{12}\,{}^{0}_{r}\lambda^{\ddagger}_{34}\right] \\ & -8\sum_{r=0,3}\sum_{\substack{1,2\\3,4}} I_{12}\left[{}^{0}_{r}P_{12}\left(\delta_{13}\delta_{24}\,\frac{1}{\Phi_{12}} - \delta_{1-2,3-4}\,\frac{2T\Gamma}{V}\right){}^{0}_{r}P^{\ddagger}_{34} - \delta_{13}\delta_{24}\,\frac{1}{\Phi_{12}}\,{}^{0}_{r}\Lambda_{12}\,{}^{0}_{r}\Lambda^{\ddagger}_{34}\right]\;. \end{aligned} \tag{60} $$

The cubic and quartic parts, respectively, of the fluctuation action have the schematic form

$$ \begin{aligned} \Delta\mathcal{A}^{(3)} \;\propto\; & G^{-3}\big[(P-\Lambda+\hat{\Phi}\hat{\gamma}q)^3 \\ & +(P-\Lambda+\hat{\Phi}\hat{\gamma}q)(q-\lambda)^2\big]\;, \end{aligned} \tag{61a} $$

$$ \begin{aligned} \Delta\mathcal{A}^{(4)} \;\propto\; & G^{-4}\big[(P-\Lambda+\hat{\Phi}\hat{\gamma}q)^4 \\ & +(P-\Lambda+\hat{\Phi}\hat{\gamma}q)^2(q-\lambda)^2 \\ & +(q-\lambda)^4\big]\;. \end{aligned} \tag{61b} $$

More explicit expression are obtained by substituting Eqs. (59) into Eqs. (53) and (54).

2. Integrating out the massive modes

We now integrate out P and Λ in a saddle-point approximation while keeping q and λ fixed. This procedure is analogous to the one explained in Appendix B for deriving the $O(2)$ nonlinear sigma model. In the current context it leads to an effective soft-mode theory that is related to, but has a different structure than, a sigma model.

The saddle-point equations read

$$ 0 \;=\; -16\,I_{12}\left[{}^{0}_{r}(\hat{\Phi}^{-1}P)^{\ddagger}_{12} - {}^{0}_{r}(\hat{\Gamma}P)^{\ddagger}_{12}\right] + \delta\Delta\mathcal{A}/\delta^{0}_{r}P_{12}\;, \tag{62a} $$

$$ 0 \;=\; 16\,I_{12}\,{}^{0}_{r}(\hat{\Phi}^{-1}\Lambda)^{\ddagger}_{12} - \delta\Delta\mathcal{A}/\delta^{0}_{r}P_{12}\;. \tag{62b} $$

Here we have used $\delta\Delta\mathcal{A}/\delta P = -\delta\Delta\mathcal{A}/\delta\Lambda$, and $\hat{\Gamma}P$ is defined in analogy to $\hat{\gamma}q$ in Eq. (59b). From Eqs. (62) we immediately obtain the useful identities

$$ \begin{aligned} (P-\Lambda)_{12} \;&=\; (\hat{\Phi}\hat{\Gamma}P)_{12} \\ &=\; (\hat{\Phi}\hat{\gamma}\Lambda)_{12}\;. \end{aligned} \tag{63} $$

P is thus given in terms of Λ, which in turn can be expressed in terms of q and λ by solving Eq. (62b). We will determine the action explicitly to $O(q^4)$,[45] which requires Λ to $O(q^2)$. From Eqs. (62b) and (61a) (or, more explicitly, from Eqs. (53) and (59); note that only $\Delta\mathcal{A}^{(3)}$ contributes to Λ to this order) we find

$$\begin{aligned}\Lambda_{12} &= -2i(\not{q}\, G^{-1}\, \not{q}^{+} + \not{q}^{+}\, G^{-1}\, \not{q})_{12} \\ &\quad -2i\,\Big[((\hat{\Phi}\hat{\gamma}q)\, G^{-1}\, (\hat{\Phi}\hat{\gamma}q^{+}))_{12} \\ &\qquad +((\hat{\Phi}\hat{\gamma}q^{+})\, G^{-1}\, (\hat{\Phi}\hat{\gamma}q))_{12} \\ &\qquad\quad +((\hat{\Phi}\hat{\gamma}q)\, G^{-1}\, (\hat{\Phi}\hat{\gamma}q))_{12}\Big] + O(q^3).\end{aligned} \tag{64}$$

Terms of higher order in q can be obtained by means of an obvious iteration procedure.

Fluctuations about this saddle point yield measure terms that reflect the Jacobian due to the change from integrating over δQ and $\delta\tilde{\Lambda}$ to obtain the partition function to integrating over q and λ. For the $O(2)$ nonlinear sigma model this is demonstrated in Appendix B. While they are structurally important for preserving the symmetry of the action, these terms are not of practical importance for a loop expansion near the lower critical dimensionality d_c^- of the problem. The reason is that they lead to loop integrals that are less infrared divergent than those that result from the terms we have derived so far. In the $O(2)$ nonlinear sigma model they carry no gradients, see Eq. (B12), and hence are less infrared divergent near $d_c^- = 2$ by two degrees than the terms contained in the nonlinear sigma model without the measure terms. The same statement is true for the generalized nonlinear sigma model that describes the Anderson-Mott metal-insulator transition in the disordered interacting fermion problem, which also has $d_c^- = 2$.[12,14] For the current problem the lower critical dimension is $d_c^- = 1$, the dimension at and below which the Fermi liquid is not stable for any parameter values. This reduction of d_c^- compared to the disordered case is a consequence of the soft modes having a ballistic dispersion relation, as opposed to a diffusive one. For the same reason the measure terms will lead to loop integrals that, near $d_c^- = 1$, are less divergent by one power of momentum than the ones we keep. For the sake of simplicity we will ignore these measure terms in what follows; it is clear how to derive them if that's desirable.

D. Effective action

1. *Effective action to $O(q^4)$*

We are now in a position to write the effective action in terms of the soft modes q and λ only. The relation between $\hat{\Gamma}P$ and $\hat{\gamma}\Lambda$ in Eq. (63) allows to write the Gaussian action, Eq. (60),

$$\mathcal{A}^{(2)} = -8\sum_{r=0,3}\sum_{\substack{1,2\\3,4}}\left[{}^{0}_{r}q_{12}\left(\delta_{13}\delta_{24}\,\frac{1}{\varphi_{12}} - \delta_{1-2,3-4}\,\frac{2T\gamma}{V}\right){}^{0}_{r}q^{\ddagger}_{34} - \delta_{13}\delta_{24}\,\frac{1}{\varphi_{12}}\,{}^{0}_{r}\lambda_{12}\,{}^{0}_{r}\lambda^{\ddagger}_{34} + {}^{0}_{r}\Lambda_{12}\,\frac{2T\gamma}{V}\,\delta_{1-2,3-4}\,{}^{0}_{r}\Lambda^{\ddagger}_{34}\right]. \tag{65}$$

This can be written more compactly

$$\begin{aligned}\mathcal{A}^{(2)} &= -2\,\mathrm{Tr}\,[q(\hat{\varphi}^{-1} - \hat{\gamma})q^{+}] + 2\,\mathrm{Tr}\,(\lambda\,\hat{\varphi}^{-1}\,\lambda^{+}) \\ &\quad -2\,\mathrm{Tr}\,(\Lambda\,\hat{\gamma}\,\Lambda^{+})\,.\end{aligned} \tag{66}$$

Here Λ is given in terms of q and λ by Eq. (64). Similarly, the other terms in the action, $\Delta\mathcal{A}^{(3)}$, $\Delta\mathcal{A}^{(4)}$, etc., can be expressed in terms of q and λ by means of Eqs. (59), (63), and (64). The resulting action is the desired effective action in terms of the soft modes q and λ.

Note that this is a local field theory, i.e., all vertices are finite in the limit of small wave numbers and small frequencies. This can be seen most easily from Eqs. (61) and their generalizations: The terms with the highest power of Green functions in $\Delta\mathcal{A}^{(n)}$ has the structure $G^{-n}\Phi^{n}(\gamma q)^{n} \sim G^{n}(\gamma q)^{n}$. Since $G \sim 1/\Omega$ and $\gamma q \sim \Omega q$, this vertex is of $O(1)$, and all other vertices scale as a positive power of the frequency.

One might consider analyzing this theory by means of a loop expansion in terms of both q and λ. However, this is not desirable. The reason is that the origin of λ is a hard (i.e., delta-function) constraint that constrains the q to soft bilinear fermion modes. Treating λ perturbatively relaxes this hard constraint, which effectively introduces spurious soft modes. (We note that no such difficulty arises for Λ, which enforces a constraint for massive modes.) It therefore is desirable to treat λ exactly. This can be achieved by eliminating λ in favor of diagram rules for an action that is

formulated entirely in terms of q and $\not{q}$. To see this let us consider the terms quadratic in q and λ, i.e., the first two terms in Eq. (65) or (66). The Gaussian q and λ-propagators are obtained by inverting the corresponding vertices. We see that the λ-propagator is minus the noninteracting part of the q-propagator, which is given by φ, and that q and λ are not coupled in the Gaussian action. This implies that whenever $\not{q}$ is contracted with $\not{q}$, the resulting propagator is effectively equal to the interacting part of the q-propagator, whereas contracting $\not{q}$ with λ results in the full q-propagator, as does contracting q with q.[46] λ can thus formally be integrated out, despite the fact that it is a soft fluctuation, without causing undesirable features of the effective theory, such as nonlocality. The reason is that the *only* effect of λ is to cancel well-defined contributions from other soft fluctuations. We thus arrive at the following effective soft-mode action:

$$\mathcal{A}_{\rm eff} = \mathcal{A}^{(2)}_{\rm eff} + \Delta\mathcal{A}^{(3)}_{\rm eff} + \Delta\mathcal{A}^{(4)}_{\rm eff} + O(q^5) \ . \tag{67a}$$

with a Gaussian part

$$\begin{aligned} \mathcal{A}^{(2)}_{\rm eff} &= -2\,{\rm Tr}\,[q(\hat\varphi^{-1}-\hat\gamma)q^+] \\ &\equiv -8\sum_{r=0,3}\sum_{\substack{1,2\\3,4}} {}^0_r q_{12}\left(\delta_{13}\delta_{24}\,\frac{1}{\varphi_{12}} - \delta_{1-2,3-4}\,\frac{2T\gamma}{V}\right){}^0_r q^\ddagger_{34} \end{aligned} \tag{67b}$$

and non-Gaussian parts $\Delta\mathcal{A}^{(n)}_{\rm eff}$ that each contain terms of $O(q^n)$. The first two we obtain from Eqs. (53), (54), and (59b) as follows,[47]

$$\begin{aligned} \Delta\mathcal{A}^{(3)}_{\rm eff} &= -4i\,{\rm Tr}\,\Big[G^{-1}\,(\hat\Phi\hat\gamma q+\hat\Phi\hat\gamma q^+)\,G^{-1}\,(\not{q}\,G^{-1}\not{q}^+ + \not{q}^+\,G^{-1}\not{q})\Big] \\ &\quad -4i\,{\rm Tr}\,\Big[G^{-1}(\hat\Phi\hat\gamma\, q)\,G^{-1}\,(\hat\Phi\hat\gamma q+\hat\Phi\hat\gamma q^+)\,G^{-1}\,(\hat\Phi\hat\gamma\, q^+)\Big]\ , \end{aligned} \tag{67c}$$

$$\begin{aligned} \Delta\mathcal{A}^{(4)}_{\rm eff} = &-4\,{\rm Tr}\,\big[G^{-1}\not{q}\,G^{-1}\not{q}^+\,G^{-1}\not{q}\,G^{-1}\not{q}^+\big] \\ &-4i\,{\rm Tr}\,\Big[G^{-1}\,(\hat\Phi\hat\gamma\Lambda+\hat\Phi\hat\gamma\Lambda^+)\,G^{-1}\,(\not{q}\,G^{-1}\not{q}^+ + \not{q}^+\,G^{-1}\not{q})\Big] \\ &-2\,{\rm Tr}\,(\Lambda\hat\gamma\Lambda^+) \\ &-8\,{\rm Tr}\,\Big[G^{-1}(\hat\Phi\hat\gamma q+\hat\Phi\hat\gamma q^+)\,G^{-1}\,(\hat\Phi\hat\gamma q+\hat\Phi\hat\gamma q^+)\,G^{-1}\,(\not{q}\,G^{-1}\not{q}^+ + \not{q}^+\,G^{-1}\not{q}) \\ &\qquad\qquad +G^{-1}\,(\hat\Phi\hat\gamma q+\hat\Phi\hat\gamma q^+)\,G^{-1}\not{q}\,G^{-1}\,(\hat\Phi\hat\gamma q+\hat\Phi\hat\gamma q^+)\,G^{-1}\not{q}^+\Big] \\ &-4i\,{\rm Tr}\,\Big[G^{-1}\,(\hat\Phi\hat\gamma\Lambda)\,G^{-1}\,(\hat\Phi\hat\gamma q+\hat\Phi\hat\gamma q^+)\,G^{-1}\,(\hat\Phi\hat\gamma q^+) + G^{-1}\,(\hat\Phi\hat\gamma q)\,G^{-1}\,(\hat\Phi\hat\gamma\Lambda+\hat\Phi\hat\gamma\Lambda^+)\,G^{-1}\,(\hat\Phi\hat\gamma q^+) \\ &\quad +G^{-1}\,(\hat\Phi\hat\gamma q)\,G^{-1}\,(\hat\Phi\hat\gamma q+\hat\Phi\hat\gamma q^+)\,G^{-1}\,(\hat\Phi\gamma\Lambda^+)\Big] \\ &-2\,{\rm Tr}\,\Big[4\,G^{-1}\,(\hat\Phi\hat\gamma q)\,G^{-1}\,(\hat\Phi\hat\gamma q)\,G^{-1}\,(\hat\Phi\hat\gamma q)\,G^{-1}\,(\hat\Phi\hat\gamma q^+) + 4\,G^{-1}\,(\hat\Phi\hat\gamma q^+)\,G^{-1}\,(\hat\Phi\hat\gamma q^+)\,G^{-1}\,(\hat\Phi\hat\gamma q^+)\,G^{-1}\,(\hat\Phi\hat\gamma q) \\ &\quad +3\,G^{-1}\,(\hat\Phi\hat\gamma q^+)\,G^{-1}\,(\hat\Phi\hat\gamma q^+)\,G^{-1}\,(\hat\Phi\hat\gamma q)\,G^{-1}\,(\hat\Phi\hat\gamma q) \\ &\quad +3\,G^{-1}\,(\hat\Phi\hat\gamma q^+)\,G^{-1}\,(\hat\Phi\hat\gamma q)\,G^{-1}\,(\hat\Phi\hat\gamma q^+)\,G^{-1}\,(\hat\Phi\hat\gamma q)\Big]\ . \end{aligned} \tag{67d}$$

Here Λ is given by Eq. (64), which we restate for completeness:

$$\begin{aligned} \Lambda &= -2i(\not{q}\,G^{-1}\not{q}^+ + \not{q}^+\,G^{-1}\not{q}) \\ &\quad -2i\left[(\hat\Phi\hat\gamma q)\,G^{-1}\,(\hat\Phi\hat\gamma q^+) + (\hat\Phi\hat\gamma q^+)\,G^{-1}\,(\hat\Phi\hat\gamma q) + (\hat\Phi\hat\gamma q^+)\,G^{-1}\,(\hat\Phi\hat\gamma q^+)\right]\ . \end{aligned} \tag{67e}$$

For the effective action to $O(q^4)$, Λ is needed only to $O(q^2)$.

This action needs to be augmented by a prescription for the role of $\not{q}$. From the above discussion it follows that contractions of $\not{q}$ with q are given by the q-propagator as it follows from $\mathcal{A}^{(2)}_{\rm eff}$,

$$\begin{aligned}\langle {}^0_r q_{12}\, {}^0_s q^{\ddagger}_{34}\rangle \;&=\; \langle {}^0_r \not{q}_{12}\, {}^0_s q^{\ddagger}_{34}\rangle = \langle {}^0_r q_{12}\, {}^0_s \not{q}^{\ddagger}_{34}\rangle = \\ &= \frac{1}{16}\,\delta_{rs}\left[\delta_{13}\delta_{24}\,\varphi_{12} + \frac{2\gamma T}{V}\delta_{1-2,3-4}\,\frac{\varphi_{12}\,\varphi_{34}}{1+2\gamma\chi^{(0)}_{1-2}}\right] ,\end{aligned} \tag{68a}$$

while contractions of $\not{q}$ with $\not{q}$ are given by

$$\langle {}^0_r \not{q}_{12}\, {}^0_s \not{q}^{\ddagger}_{34}\rangle = \frac{\gamma T}{8V}\,\delta_{1-2,3-4}\,\frac{\varphi_{12}\,\varphi_{34}}{1+2\gamma\chi^{(0)}_{1-2}}\ . \tag{68b}$$

Here we have defined

$$\chi^{(0)}_{1-2} = -\frac{T}{V}\sum_{34}\delta_{1-2,3-4}\,\varphi_{34}\ . \tag{69}$$

Physically, $\chi^{(0)}$ is the hydrodynamic part of the density susceptibility per spin in Hartree-Fock approximation, see the discussion below.

Equations (67) through (69) completely specify the effective theory for the purposes of a loop expansion for q-correlation functions.[48] They are the central formal result of the present paper. Their derivation makes is clear how to derive the effective action to any desired power in q.

2. Gaussian propagators

In order to make the theory suitable for actual calculations we need to explicitly evaluate the correlation function $\chi^{(0)}$ defined in Eq. (69). If we absorb the Hartree-Fock self energy into the chemical potential, as explained in Sec. IV B 1, the integral that defines $\chi^{(0)}$ reads explicitly

$$\begin{aligned}\chi^{(0)}(\boldsymbol{k}, i\Omega_n) \;&=\; -T\sum_{m=0}^{n-1}\frac{1}{V}\sum_{\boldsymbol{p}}\frac{1}{i\omega_m - \xi_{\boldsymbol{p}+\boldsymbol{k}/2}} \\ &\quad\times\frac{1}{i\omega_m - i\Omega_n - \xi_{\boldsymbol{p}-\boldsymbol{k}/2}}\ .\end{aligned} \tag{70}$$

The integral can be performed exactly. However, this is not necessary for most purposes, as the crucial hydrodynamic structure of $\chi^{(0)}$ is preserved in the well-known approximation, valid for small values of $k \equiv |\boldsymbol{k}|$ and $|\Omega_n|$, that performs the radial part of the momentum integral by means of a contour integration over the interval $-\infty < \xi_{\boldsymbol{p}} < \infty$, which makes the frequency sum trivial.[3] Within this scheme, which we refer to as the AGD approximation, we are left with an angular integral only,

$$\chi^{(0)}(\boldsymbol{k}, i\Omega_n) = -N_{\rm F}\,\frac{G\Omega_n}{k}\,\varphi_d(Gi\Omega_n/k)\ , \tag{71a}$$

where

$$\varphi_d(z) = \frac{-i}{S_{d-1}}\int d\Omega_{\boldsymbol{p}}\,\frac{1}{\hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{k}} - z} \tag{71b}$$

is a causal function of the complex frequency z. G is a coupling constant whose bare value is

$$G = 1/v_{\rm F}\ . \tag{71c}$$

It constitutes the natural coupling constant for a loop expansion. $S_{d-1} = 2\pi^{d/2}/\Gamma(d/2)$ with $\Gamma(x)$ the Gamma function is the surface area of the unit $(d-1)$-sphere, and $d\Omega_{\boldsymbol{p}}$ is the angular integration measure for the unit vector $\hat{\boldsymbol{p}}$ for fixed unit vector $\hat{\boldsymbol{k}}$ in d dimensions. For general dimensions $d > 1$ the latter takes the form[49,50]

$$\int d\Omega_{\boldsymbol{p}} = \frac{2\pi^{\epsilon/2}}{\Gamma(\epsilon/2)}\int_{-1}^{1}\frac{d\eta}{(1-\eta^2)^{1-\epsilon/2}}\ ,\ (\epsilon = d-1 > 0)\ , \tag{72a}$$

where $\epsilon = d-1$ and $\eta = \cos\theta$ with $\theta = \sphericalangle(\hat{\boldsymbol{p}}, \hat{\boldsymbol{k}})$. The limiting expression for $d \to 1$ is

$$\int d\Omega_{\boldsymbol{p}} = \int_{-\infty}^{\infty} d\eta \; [\delta(\eta+1) + \delta(\eta-1)] \quad , \quad (d=1) \; . \tag{72b}$$

From Eqs. (71b), (72) we obtain

$$\varphi_d(z) = \frac{i}{z} \, {}_2F_1(1, 1/2, d/2; 1/z^2) \; , \tag{73}$$

where ${}_2F_1$ is Gauss's hypergeometric function. In $d = 1, 2, 3$ this reduces to the familiar expressions for the hydrodynamic part of the Lindhard function in these dimensions:

$$\varphi_{d=1}(z) \;=\; -iz/(1-z^2) \; , \tag{74a}$$

$$\begin{aligned} \varphi_{d=2}(z) \;&=\; \mathrm{sgn}\,(\mathrm{Im}\, z)/\sqrt{1-z^2} \\ &\equiv\; i/\sqrt{z+1}\sqrt{z-1} \; , \end{aligned} \tag{74b}$$

$$\varphi_{d=3}(z) \;=\; \frac{-i}{2} \, \ln\left(\frac{1-z}{-1-z}\right) \; . \tag{74c}$$

The Gaussian propagators, Eqs. (68), are now explicitly specified.

3. The Goldstone propagator and the density susceptibility

From the preceding two subsections we see that the Goldstone modes are given by the propagator $\varphi_d(z)$, Eq. (73), which determines the density susceptibility in saddle-point (i.e., Hartree-Fock) approximation via Eq. (71a). This is just the hydrodynamic part of the familiar Lindhard function, see Eqs. (74). It is illustrative to also consider the density susceptibility in the Gaussian approximation to the effective field theory. From Eq. (18a) we see that the density susceptibility χ is given by a $\langle q\, q\rangle$ correlation function plus terms that result from the coupling of q to $\bar{P}$. Eliminating that coupling by the shift that was discussed in Sec. IV C 1 leads to Fermi-liquid corrections. Using Eq. (68a) we obtain for the density susceptibility in Gaussian approximation

$$\chi(\boldsymbol{k}, i\Omega_n) = \frac{1}{(1+2N_{\mathrm{F}}\Gamma)^2} \, \frac{2\chi^{(0)}(\boldsymbol{k}, i\Omega_n)}{1 + 2\gamma\chi^{(0)}(\boldsymbol{k}, i\Omega_n)} \; . \tag{75}$$

We recognize this as the random-phase approximation (RPA) for the density susceptibility[51] with additional Fermi-liquid corrections. One of its characteristic features is the collective mode known as zero sound that results from a real root of the denominator. To see this we define another causal function

$$\varphi_d^{(\mathrm{s})}(z) = \frac{1}{\varphi_d^{-1}(z) + i2N_{\mathrm{F}}\gamma z} \; . \tag{76}$$

χ can be expressed in terms of this as

$$\chi(\boldsymbol{k}, i\Omega_n) = \frac{-2N_{\mathrm{F}}}{(1+2N_{\mathrm{F}}\Gamma)^2} \, \frac{G\Omega_n}{k} \, \varphi_d^{(\mathrm{s})}(Gi\Omega_n/k) \; . \tag{77}$$

From Eq. (71a) we see that, apart from the Fermi-liquid corrections, $\varphi_d^{(\mathrm{s})}$ relates to φ_d the same way χ relates to $\chi^{(0)}$. Using Eqs. (73, 74) in Eq. (76) one finds that in $d = 1$ the spectrum of $\varphi_d^{(\mathrm{s})}$ is exhausted by the zero-sound poles. In $d > 1$ a particle-hole continuum emerges, but the zero-sound poles outside of the continuum remain. This is illustrated in Fig. 1. Note that the zero-sound modes are collective density fluctuations that are the result of particle-number conservation, whereas the Goldstone modes result from the spontaneously broken continuous symmetry discussed in Sec. III and are *not* related to a conservation law. This manifests itself, for instance, in the fact that the Goldstone modes acquire a mass at nonzero temperatures, whereas the zero-sound mode does not.

Equation (77) provides an important identity that allows to identify the thermodynamic density susceptibility $\partial n/\partial\mu$ in terms of the parameters of the field theory. For simplicity we consider the case $d = 1$. Using Eqs. (74a) and (76) in (77) yields

$$\begin{aligned} \chi(\boldsymbol{k}, i\Omega_n) \;&=\; \frac{-2N_{\mathrm{F}}}{1+2N_{\mathrm{F}}\Gamma} \, \frac{\Omega^2}{(1+2N_{\mathrm{F}}\Gamma)k^2/G^2 + \Omega^2} \\ &\qquad\qquad (d=1) \; . \end{aligned} \tag{78a}$$

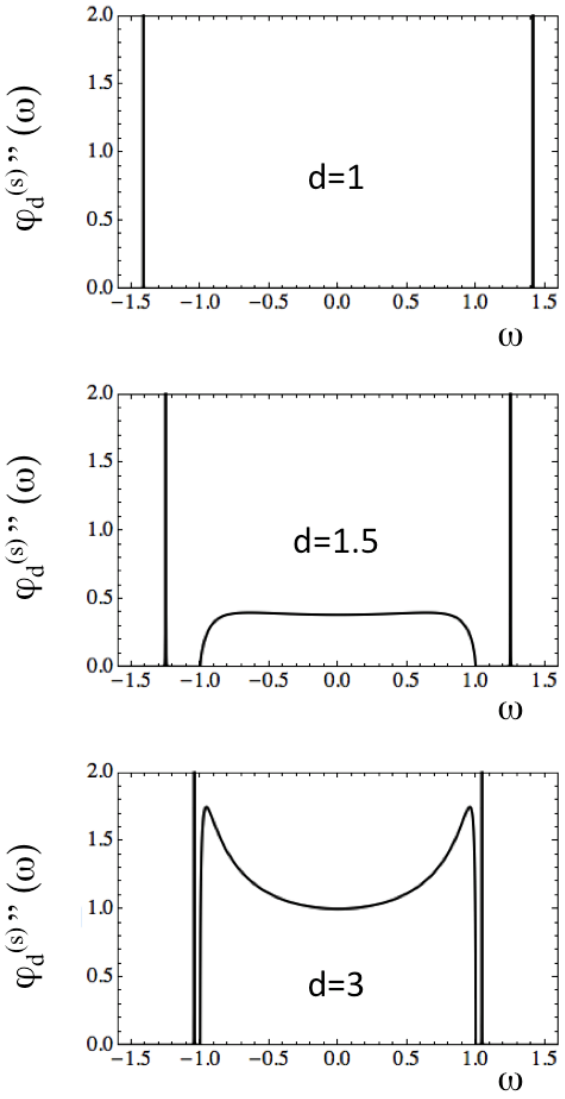

FIG. 1: The spectrum $\varphi_d^{(\mathrm{s})\prime\prime}(\omega) = \mathrm{Re}\,\varphi_d^{(\mathrm{s})}(z = \omega + i0)$ for $2N_\mathrm{F}\gamma = 0.5$ as a function of the frequency ω in dimensions $d = 1$, $d = 1.5$, and $d = 3$. The vertical lines denote δ-function contributions to the spectrum.

The compressibility sum rule,[51]

$$\partial n/\partial\mu = -\chi(\boldsymbol{k} = 0, i\Omega_n) \tag{78b}$$

then yields

$$\frac{\partial n}{\partial \mu} = \frac{2N_\mathrm{F}}{1 + 2N_\mathrm{F}\Gamma} \,. \tag{79}$$

This is consistent with another argument. $\partial n/\partial\mu$ is related to the zero-sound speed s by $s^2 = (n/m)/(\partial n/\partial\mu)$.[3] From Eq. (78a) we see that $s^2 = (1 + 2N_\mathrm{F}\Gamma)/G^2$, which also yields Eq. (79).

While we have derived this relation from the Gaussian theory, it is expected to be an exact identity that remains true if $\partial n/\partial\mu$ and Γ are replaced by their renormalized counterparts. This is because the structure of the density susceptibility, which is governed by particle-number conservation, must stay the same. This is an important point since it is known from perturbation theory that $\partial n/\partial\mu$ is not singularly renormalized near $d = 1$.[52]

E. Density formulation of the effective field theory

As we have mentioned after Eq. (50) and (66), the current theory explicitly keeps all of the soft modes that were identified by the Ward identity in Sec. III and hence is local. This is in contrast to the theory formulated in Ref. 16, which was a formulation in terms of s-wave ($\ell = 0$) modes only. Technically this is achieved by constraining $B_{nm}(\boldsymbol{x}, \boldsymbol{x})$, whose spatial Fourier transform is given by Eq. (10a), to a local matrix $Q_{nm}(\boldsymbol{x})$ by means of a Lagrange multiplier field $\tilde{\Lambda}_{nm}(\boldsymbol{x})$, and formulating the theory in terms of $Q(\boldsymbol{x})$ and $\tilde{\Lambda}(\boldsymbol{x})$. This amounts to integrating out the modes in the higher angular momentum channels, and since these are soft the result is a nonlocal field theory. While this is undesirable for many purposes, for certain calculations it may be advantageous since it makes it easier to see which loop integrals as infrared divergent and which are not. The nonlocality, on the other hand, makes no difference in practice unless one is explicitly interested in angular momentum channels with $\ell \geq 1$ or wishes to calculate non-local correlation functions. In the remainder of the section we therefore present the theory in such a density formulation.

The Fermi-liquid saddle-point solution is the same as in Sec. IV A 1, except that the saddle-point value of Q is a position-independent object Q_n given by Eq. (38a) summed over the wave vector $\boldsymbol{k}$. The further development proceeds as in the phase-space formulation, except that the various products of Green functions or their inverses get replaced by products of convolutions of Green functions in wave-vector space and inverses of such convolutions. The frequency structure of the theory remains unchanged. The Gaussian action, after decoupling $\bar{P}$ and q, then reads

$$\begin{aligned}\mathcal{A}^{(2)} \;=\; & -8 \sum_{r=0,3} \sum_{\substack{1,2\\3,4}} \frac{1}{V} \sum_{\boldsymbol{k}} \left[{}^{0}_{r}q_{12}(\boldsymbol{k}) \left(\delta_{13}\,\delta_{24}\,\frac{1}{\varphi_{12}(\boldsymbol{k})} - \delta_{1-2,3-4} 2T\gamma \right) {}^{0}_{r}q_{34}(-\boldsymbol{k}) - \delta_{13}\,\delta_{24}\,\frac{1}{\varphi_{12}(\boldsymbol{k})}\, {}^{0}_{r}\lambda_{12}(\boldsymbol{k}) {}^{0}_{r}\lambda_{34}(-\boldsymbol{k}) \right] \\ & -8 \sum_{r=0,3} \sum_{\substack{1,2\\3,4}} I_{12} \frac{1}{V} \sum_{\boldsymbol{k}} \left[{}^{0}_{r}P_{12}(\boldsymbol{k}) \left(\delta_{13}\,\delta_{24}\,\frac{1}{\Phi_{12}(\boldsymbol{k})} - \delta_{1-2,3-4} 2T\Gamma \right) {}^{0}_{r}P_{34}(-\boldsymbol{k}) - \delta_{13}\,\delta_{24}\,\frac{1}{\Phi_{12}(\boldsymbol{k})}\, {}^{0}_{r}\Lambda_{12}(\boldsymbol{k}) {}^{0}_{r}\Lambda_{34}(-\boldsymbol{k}) \right] .\end{aligned} \tag{80}$$

This replaces Eq. (60). In this subsection, $1 \equiv n_1$, etc. represents Matusbara frequency labels only, in contrast to the notation used in the other parts of this paper. We have defined

$$\begin{aligned}\phi_{12}(\boldsymbol{k}) \;&=\; \frac{1}{V} \sum_{\boldsymbol{p}} G_{n_1}(\boldsymbol{p})\, G_{n_2}(\boldsymbol{p}-\boldsymbol{k}) \\ &\equiv\; \begin{cases} \Phi_{12}(\boldsymbol{k}) & \text{if } (n_1+1/2)(n_2+1/2) > 0 \\ \varphi_{12}(\boldsymbol{k}) & \text{if } (n_1+1/2)(n_2+1/2) < 0 \,, \end{cases}\end{aligned} \tag{81a}$$

which replaces Eq. (47c). Similarly, the products of inverse Green functions in the higher terms in the fluctuation expansion get replaced by more complex expressions, which involve convolutions of Green functions, with the same net power of Green functions. To account for this, we define

$$\begin{aligned}\phi^{(m)}_{1\ldots m}(\boldsymbol{k}_1, \ldots, \boldsymbol{k}_m) \;&=\; \frac{1}{V} \sum_{\boldsymbol{p}} G_{n_1}(\boldsymbol{p})\, G_{n_2}(\boldsymbol{p}-\boldsymbol{k}_1-\boldsymbol{k}_2) \cdots G_{n_m}(\boldsymbol{p}-\boldsymbol{k}_1-\ldots-\boldsymbol{k}_m) \\ &\equiv\; \begin{cases} \Phi^{(m)}_{1\ldots m}(\boldsymbol{k}_1, \ldots, \boldsymbol{k}_m) & \text{if } n_1 \text{ through } n_m \text{ all have the same sign} \\ \varphi^{(m)}_{1\ldots m}(\boldsymbol{k}_1, \ldots, \boldsymbol{k}_m) & \text{if } n_1 \text{ through } n_m \text{ do not all have the same sign} \,, \end{cases}\end{aligned} \tag{81b}$$

The cubic and quartic terms, instead of Eqs. (53, 54), now take the form

$$\begin{aligned}\Delta\mathcal{A}^{(3)} \;=\; & \frac{-4i}{V^2} \sum_{1,2,3} \sum_{\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3} \delta_{\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3,0} \\ & \times \Big\{ \Phi^{(3)}_{123}(\boldsymbol{k}_1,\boldsymbol{k}_2)\, \Phi^{-1}_{12}(\boldsymbol{k}_1)\, \Phi^{-1}_{23}(\boldsymbol{k}_2)\, \Phi^{-1}_{31}(\boldsymbol{k}_3) \Big\{ \mathrm{tr}\, \left[(\bar{P}-\Lambda)_{12}(\boldsymbol{k}_1) (\bar{P}+\bar{P}^{+}-\Lambda-\Lambda^{+})_{23}(\boldsymbol{k}_2) (\bar{P}^{+}-\Lambda^{+})_{31}(\boldsymbol{k}_3) \right] \\ & + \varphi^{(3)}_{123}(\boldsymbol{k}_1,\boldsymbol{k}_2)\, \Phi^{-1}_{12}(\boldsymbol{k}_1)\, \varphi^{-1}_{23}(\boldsymbol{k}_2)\, \varphi^{-1}_{31}(\boldsymbol{k}_3)\, \mathrm{tr}\, \left[(\bar{P}+\bar{P}^{+}-\Lambda-\Lambda^{+})_{12}(\boldsymbol{k}_1) \big(\not{q}_{23}(\boldsymbol{k}_2) \not{q}^{+}_{31}(\boldsymbol{k}_3) + \not{q}^{+}_{23}(\boldsymbol{k}_2) \not{q}_{31}(\boldsymbol{k}_3) \big) \right] \Big\} .\end{aligned} \tag{82a}$$

$$\Delta\mathcal{A}^{(4)} = \Delta\mathcal{A}^{(4,0)} + \Delta\mathcal{A}^{(2,2)} + \Delta\mathcal{A}^{(0,4)} \,, \tag{82b}$$

with

$$\begin{aligned}\Delta\mathcal{A}^{(4,0)} &= \frac{-2}{V^3}\sum_{1,2,3,4}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4}\delta_{\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3+\boldsymbol{k}_4,0}\,\Phi^{(4)}_{1234}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)\,\Phi^{-1}_{12}(\boldsymbol{k}_1)\,\Phi^{-1}_{23}(\boldsymbol{k}_2)\,\Phi^{-1}_{34}(\boldsymbol{k}_3)\,\Phi^{-1}_{41}(\boldsymbol{k}_4)\\ &\quad\times\Big[4\,\mathrm{tr}\,\big[(\bar P-\Lambda)_{12}(\boldsymbol{k}_1)(\bar P-\Lambda)_{23}(\boldsymbol{k}_2)(\bar P-\Lambda)_{34}(\boldsymbol{k}_3)(\bar P-\Lambda)^+_{41}(\boldsymbol{k}_4)\big]\\ &\quad\;+4\,\mathrm{tr}\,\big[(\bar P-\Lambda)^+_{12}(\boldsymbol{k}_1)(\bar P-\Lambda)^+_{23}(\boldsymbol{k}_2)(\bar P-\Lambda)^+_{34}(\boldsymbol{k}_3)(\bar P-\Lambda)_{41}(\boldsymbol{k}_4)\big]\\ &\quad\;+3\,\mathrm{tr}\,\big[(\bar P-\Lambda)_{12}(\boldsymbol{k}_1)(\bar P-\Lambda)_{23}(\boldsymbol{k}_2)(\bar P-\Lambda)^+_{34}(\boldsymbol{k}_3)(\bar P-\Lambda)^+_{41}(\boldsymbol{k}_4)\big]\\ &\quad\;+3\,\mathrm{tr}\,\big[(\bar P-\Lambda)_{12}(\boldsymbol{k}_1)(\bar P-\Lambda)^+_{23}(\boldsymbol{k}_2)(\bar P-\Lambda)_{34}(\boldsymbol{k}_3)(\bar P-\Lambda)^+_{41}(\boldsymbol{k}_4)\big]\Big]\,,\end{aligned}\tag{82c}$$

$$\begin{aligned}\Delta\mathcal{A}^{(2,2)} &= \frac{-8}{V^3}\sum_{1,2,3,4}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4}\delta_{\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3+\boldsymbol{k}_4,0}\,\varphi^{(4)}_{1234}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)\\ &\quad\times\Big\{\Phi^{-1}_{12}(\boldsymbol{k}_1)\,\Phi^{-1}_{23}(\boldsymbol{k}_2)\,\varphi^{-1}_{34}(\boldsymbol{k}_3)\,\varphi^{-1}_{41}(\boldsymbol{k}_4)\,\mathrm{tr}\,\big[(\bar P+\bar P^+-\Lambda-\Lambda^+)_{12}(\boldsymbol{k}_1)(\bar P+\bar P^+-\Lambda-\Lambda^+)_{23}(\boldsymbol{k}_2)\\ &\qquad\qquad\times\big(\not{g}_{34}(\boldsymbol{k}_3)\,\not{g}^+_{41}(\boldsymbol{k}_4)+\not{g}^+_{34}(\boldsymbol{k}_3)\,\not{g}_{41}(\boldsymbol{k}_4)\big)\big]\\ &\quad\;+\Phi^{-1}_{12}(\boldsymbol{k}_1)\,\varphi^{-1}_{23}(\boldsymbol{k}_2)\,\Phi^{-1}_{34}(\boldsymbol{k}_3)\,\varphi^{-1}_{41}(\boldsymbol{k}_4)\,\mathrm{tr}\,\big[(\bar P+\bar P^+-\Lambda-\Lambda^+)_{12}(\boldsymbol{k}_1)\,\not{g}_{23}(\boldsymbol{k}_2)\\ &\qquad\qquad\times(\bar P+\bar P^+-\Lambda-\Lambda^+)_{34}(\boldsymbol{k}_3)\,\not{g}^+_{41}(\boldsymbol{k}_4)\big]\Big\}\,,\end{aligned}\tag{82d}$$

$$\begin{aligned}\Delta\mathcal{A}^{(0,4)} &= \frac{-4}{V^3}\sum_{1,2,3,4}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4}\delta_{\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3+\boldsymbol{k}_4,0}\,\varphi^{(4)}_{1234}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3)\,\varphi^{-1}_{12}(\boldsymbol{k}_1)\,\varphi^{-1}_{23}(\boldsymbol{k}_2)\,\varphi^{-1}_{34}(\boldsymbol{k}_3)\,\varphi^{-1}_{41}(\boldsymbol{k}_4)\\ &\quad\times\mathrm{tr}\,\big[\not{g}_{12}(\boldsymbol{k}_1)\,\not{g}^+_{23}(\boldsymbol{k}_2)\,\not{g}_{34}(\boldsymbol{k}_3)\,\not{g}^+_{41}(\boldsymbol{k}_4)\big]\,.\end{aligned}\tag{82e}$$

The relation between $\bar P$ and P is

$$ {}^0_r P_{12}(\boldsymbol{k}) = {}^0_r\bar P_{12}(\boldsymbol{k}) - \Phi_{12}(\boldsymbol{k})\,2\gamma T\sum_{3,4}\delta_{1-2,3-4}\,{}^0_r q_{34}(\boldsymbol{k})\,,\tag{83}$$

which replaces Eq. (59a). Note that the vertices in Eqs. (82) are invariant under cyclic permutations of the frequency and wave vector indices. Also note that these vertices are in general not finite in the hydrodynamic limit of small wave vectors and frequencies, since the n-point susceptibility $\varphi^{(n)}_{12\ldots n}(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n)$ diverges in this limit.[16] This formulation therefore does not yield a local field theory.

P and Λ can now be eliminated in analogy to Sec. IV C 2. The identities (63) turn into

$$\begin{aligned}(P-\Lambda)_{12}(\boldsymbol{k}) &= \Phi_{12}(\boldsymbol{k})\,(\hat\Gamma P)_{12}(\boldsymbol{k})\\ &= \Phi_{12}(\boldsymbol{k})\,(\hat\gamma\Lambda)_{12}(\boldsymbol{k})\,,\end{aligned}\tag{84a}$$

with

$$(\hat\Gamma P)_{12}(\boldsymbol{k}) = 2T\Gamma\sum_{3,4}\delta_{1-2,3-4}\,P_{34}(\boldsymbol{k})\,.\tag{84b}$$

$(\hat\gamma\Lambda)$ is defined analogously, and so is $(\hat\gamma q)$. Λ in terms of q takes the form

$$\begin{aligned}\Lambda_{12}(\boldsymbol{k}) &= -2i\sum_3\frac{1}{V}\sum_{\boldsymbol{p}}\varphi^{(3)}_{132}(\boldsymbol{p},\boldsymbol{k}-\boldsymbol{p})\,\varphi^{-1}_{13}(\boldsymbol{p})\,\varphi^{-1}_{32}(\boldsymbol{k}-\boldsymbol{p})\,\big[\not{g}_{13}(\boldsymbol{p})\,\not{g}^+_{32}(\boldsymbol{k}-\boldsymbol{p})+\not{g}^+_{13}(\boldsymbol{p})\,\not{g}_{32}(\boldsymbol{k}-\boldsymbol{p})\big]\\ &\quad-2i\sum_3\frac{1}{V}\sum_{\boldsymbol{p}}\Phi^{(3)}_{132}(\boldsymbol{p},\boldsymbol{k}-\boldsymbol{p})\,\big[(\hat\gamma q)_{13}(\boldsymbol{p})\,(\hat\gamma q^+)_{32}(\boldsymbol{k}-\boldsymbol{p})+(\hat\gamma q^+)_{13}(\boldsymbol{p})(\hat\gamma q)_{32}(\boldsymbol{k}-\boldsymbol{p})\\ &\qquad\qquad+(\hat\gamma q)_{13}(\boldsymbol{p})(\hat\gamma q)_{32}(\boldsymbol{k}-\boldsymbol{p})\big]+O(q^3)\,,\end{aligned}\tag{85}$$

which replaces Eq. (64).

Note that here, and in Eqs. (86) below, $\hat\gamma q$ always arises from a $\bar P$ that was shifted by means of Eq. (83), and therefore carries the same implicit frequency restrictions as P and Λ. That is, $(\hat\gamma q)_{12}(\boldsymbol{k})$ implies $(n_1+1/2)(n_2+1/2)>0$. A

related remark is that the notation $\varphi^{(3)}_{123}$, $\varphi^{(4)}_{1234}$, etc. just specifies that the convolution has hydrodynamic content, i.e., that all of the frequencies do not have the same sign, but not which frequencies are positive or negative. This information, which is necessary for calculating physical quantities, can be obtained by an inspection of the matrix fields that multiply the convolution.

The effective action analogous to Eqs. (67) then reads

$$\mathcal{A}_{\rm eff} = \mathcal{A}^{(2)}_{\rm eff} + \Delta\mathcal{A}^{(3)}_{\rm eff} + \Delta\mathcal{A}^{(4)}_{\rm eff} + O(q^5) \ . \tag{86a}$$

with a Gaussian part

$$\mathcal{A}^{(2)}_{\rm eff} = -8 \sum_{r=0,3} \sum_{\substack{1,2\\3,4}} \frac{1}{V} \sum_{\boldsymbol{k}} {}^{0}_{r}q_{12}(\boldsymbol{k}) \left(\delta_{13}\delta_{24} \, \frac{1}{\varphi_{12}(\boldsymbol{k})} - \delta_{1-2,3-4} \, 2T\gamma \right) {}^{0}_{r}q_{34}(-\boldsymbol{k}) \tag{86b}$$

and nonlinearities

$$\begin{aligned}
\mathcal{A}^{(3)}_{\rm eff} \ = \ & \frac{-4i}{V^2} \sum_{\substack{1,2\\3}} \sum_{\substack{\boldsymbol{k}_1,\boldsymbol{k}_2\\ \boldsymbol{k}_3}} \delta_{\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3,0} \\
& \times \Big\{ \varphi^{(3)}_{123}(\boldsymbol{k}_1,\boldsymbol{k}_2) \, \varphi^{-1}_{23}(\boldsymbol{k}_2) \, \varphi^{-1}_{31}(\boldsymbol{k}_3) \mathrm{tr} \, \big[(\hat{\gamma}q + \hat{\gamma}q^+)_{12}(\boldsymbol{k}_1) \, \big(\not{q}_{23}(\boldsymbol{k}_2) \, \not{q}^+_{31}(\boldsymbol{k}_3) + \not{q}^+_{23}(\boldsymbol{k}_2) \, \not{q}_{31}(\boldsymbol{k}_3) \big) \big] \\
& \quad + \Phi^{(3)}_{123}(\boldsymbol{k}_1,\boldsymbol{k}_2) \, \mathrm{tr} \, \big[(\hat{\gamma}q)_{12}(\boldsymbol{k}_1) \, (\hat{\gamma}q + \hat{\gamma}q^+)_{23}(\boldsymbol{k}_2) \, (\hat{\gamma}q^+)_{31}(\boldsymbol{k}_3) \big] \Big\} \ ,
\end{aligned} \tag{86c}$$

$$\begin{aligned}
\mathcal{A}^{(4)}_{\rm eff} \ = \ & - \frac{4}{V^3} \sum_{\substack{1,2\\3,4}} \sum_{\substack{\boldsymbol{k}_1,\boldsymbol{k}_2\\ \boldsymbol{k}_3,\boldsymbol{k}_4}} \delta_{\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3+\boldsymbol{k}_4,0} \, \varphi^{(4)}_{1234}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) \, \varphi^{-1}_{12}(\boldsymbol{k}_1) \, \varphi^{-1}_{23}(\boldsymbol{k}_2) \, \varphi^{-1}_{34}(\boldsymbol{k}_3) \, \varphi^{-1}_{41}(\boldsymbol{k}_4) \\
& \qquad \times \mathrm{tr} \, \big[\not{q}_{12}(\boldsymbol{k}_1) \, \not{q}^+_{23}(\boldsymbol{k}_2) \, \not{q}_{34}(\boldsymbol{k}_3) \, \not{q}^+_{41}(\boldsymbol{k}_4) \big] \\
& - \frac{4i}{V^2} \sum_{\substack{1,2\\3}} \sum_{\substack{\boldsymbol{k}_1,\boldsymbol{k}_2\\ \boldsymbol{k}_3}} \delta_{\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3,0} \, \varphi^{(3)}_{123}(\boldsymbol{k}_1,\boldsymbol{k}_2) \, \varphi^{-1}_{23}(\boldsymbol{k}_2) \, \varphi^{-1}_{31}(\boldsymbol{k}_3) \\
& \qquad \times \mathrm{tr} \, \big[(\gamma\Lambda + \gamma\Lambda^+)_{12}(\boldsymbol{k}_1) \, \big(\not{q}_{23}(\boldsymbol{k}_2) \, \not{q}^+_{31}(\boldsymbol{k}_3) + \not{q}^+(\boldsymbol{k}_2) \, \not{q}_{31}(\boldsymbol{k}_3) \big) \big] \\
& - 2 \sum_{12} \frac{1}{V} \sum_{\boldsymbol{k}} \mathrm{tr} \, \big[\Lambda_{12}(\boldsymbol{k}) \, \hat{\gamma} \Lambda^+_{21}(\boldsymbol{k}) \big] \\
& - \frac{8}{V^3} \sum_{\substack{1,2\\3,4}} \sum_{\substack{\boldsymbol{k}_1,\boldsymbol{k}_2\\ \boldsymbol{k}_3,\boldsymbol{k}_4}} \delta_{\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3+\boldsymbol{k}_4,0} \, \varphi^{(4)}_{1234}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) \\
& \qquad \times \Big\{ \varphi^{-1}_{34}(\boldsymbol{k}_3) \, \varphi^{-1}_{41}(\boldsymbol{k}_4) \, \mathrm{tr} \, \big[(\hat{\gamma}q + \hat{\gamma}q^+)_{12}(\boldsymbol{k}_1) \, (\hat{\gamma}q + \hat{\gamma}q^+)_{23}(\boldsymbol{k}_2) \, \big(\not{q}_{34}(\boldsymbol{k}_3) \not{q}^+_{41}(\boldsymbol{k}_4) + \not{q}^+_{34}(\boldsymbol{k}_3) \, \not{q}_{41}(\boldsymbol{k}_4) \big) \big] \\
& \qquad \quad + \varphi^{-1}_{23}(\boldsymbol{k}_2) \, \varphi^{-1}_{41}(\boldsymbol{k}_4) \, \mathrm{tr} \, \big[(\hat{\gamma}q + \hat{\gamma}q^+)_{12}(\boldsymbol{k}_1) \, \not{q}_{23}(\boldsymbol{k}_2) \, (\hat{\gamma}q + \hat{\gamma}q^+)_{34}(\boldsymbol{k}_3) \, \not{q}^+_{41}(\boldsymbol{k}_4) \big] \Big\} \\
& - \frac{4i}{V^2} \sum_{\substack{1,2\\3}} \sum_{\substack{\boldsymbol{k}_1,\boldsymbol{k}_2\\ \boldsymbol{k}_3}} \delta_{\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3,0} \, \Phi^{(3)}_{123}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) \Big\{ \mathrm{tr} \, \big[(\hat{\gamma}\Lambda)_{12}(\boldsymbol{k}_1) \, (\hat{\gamma}q + \hat{\gamma}q^+)_{23}(\boldsymbol{k}_2) \, (\hat{\gamma}q^+)_{31}(\boldsymbol{k}_3) \big] \\
& \qquad \qquad + \mathrm{tr} \, \big[(\gamma q)_{12}(\boldsymbol{k}_1) \, (\hat{\gamma}\Lambda + \hat{\gamma}\Lambda^+)_{23}(\boldsymbol{k}_2) \, (\hat{\gamma}q^+)_{31}(\boldsymbol{k}_3) \big] \\
& \qquad \qquad + \mathrm{tr} \, \big[(\hat{\gamma}q)_{12}(\boldsymbol{k}_1) \, (\hat{\gamma}q + \hat{\gamma}q^+)_{23}(\boldsymbol{k}_2) \, (\hat{\gamma}\Lambda^+)_{31}(\boldsymbol{k}_3) \big] \Big\} \\
& - \frac{2}{V^3} \sum_{\substack{1,2\\3,4}} \sum_{\substack{\boldsymbol{k}_1,\boldsymbol{k}_2\\ \boldsymbol{k}_3,\boldsymbol{k}_4}} \delta_{\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3+\boldsymbol{k}_4,0} \, \Phi^{(4)}_{1234}(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) \Big\{ 4 \, \mathrm{tr} \, \big[(\hat{\gamma}q)_{12}(\boldsymbol{k}_1) \, (\hat{\gamma}q)_{23}(\boldsymbol{k}_2 \, (\hat{\gamma}q)_{34}(\boldsymbol{k}_3) \, (\hat{\gamma}q^+)_{41}(\boldsymbol{k}_4) \big] \\
& \qquad \qquad + 4 \, \mathrm{tr} \, \big[(\hat{\gamma}q^+)_{12}(\boldsymbol{k}_1) \, (\hat{\gamma}q^+)_{23}(\boldsymbol{k}_2 \, (\hat{\gamma}q^+)_{34}(\boldsymbol{k}_3) \, (\hat{\gamma}q)_{41}(\boldsymbol{k}_4) \big] \\
& \qquad \qquad + 3 \, \mathrm{tr} \, \big[(\hat{\gamma}q)_{12}(\boldsymbol{k}_1) \, (\hat{\gamma}q)_{23}(\boldsymbol{k}_2 \, (\hat{\gamma}q^+)_{34}(\boldsymbol{k}_3) \, (\hat{\gamma}q^+)_{41}(\boldsymbol{k}_4) \big] \\
& \qquad \qquad + 3 \, \mathrm{tr} \, \big[(\hat{\gamma}q)_{12}(\boldsymbol{k}_1) \, (\hat{\gamma}q^+)_{23}(\boldsymbol{k}_2 \, (\hat{\gamma}q)_{34}(\boldsymbol{k}_3) \, (\hat{\gamma}q^+)_{41}(\boldsymbol{k}_4) \big] \Big\} \ ,
\end{aligned} \tag{86d}$$

where Λ in terms of q is given by Eq. (85). The rules about treating $\not{q}$ are the same as in the phase space formulation,

and the Gaussian contractions are

$$\begin{aligned}\langle {}^0_r q_{12}(\boldsymbol{k})\, {}^0_s q_{34}(-\boldsymbol{k})\rangle &= \langle {}^0_r q_{12}(\boldsymbol{k})\, {}^0_s \not{q}_{34}(-\boldsymbol{k})\rangle = \langle {}^0_r \not{q}_{12}(\boldsymbol{k})\, {}^0_s q_{34}(-\boldsymbol{k})\rangle \\ &= \frac{V}{16}\,\delta_{rs}\left[\delta_{13}\,\delta_{24}\,\varphi_{12}(\boldsymbol{k}) + 2\gamma T\,\delta_{1-2,3-4}\,\frac{\varphi_{12}(\boldsymbol{k})\,\varphi_{34}(\boldsymbol{k})}{1+2\gamma\chi^{(0)}_{1-2}(\boldsymbol{k})}\right]\,, \end{aligned} \tag{87a}$$

$$\langle {}^0_r \not{q}_{12}(\boldsymbol{k})\, {}^0_s \not{q}_{34}(-\boldsymbol{k})\rangle = \frac{V}{8}\,\gamma T\,\delta_{1-2,3-4}\,\frac{\varphi_{12}(\boldsymbol{k})\,\varphi_{34}(\boldsymbol{k})}{1+2\gamma\chi^{(0)}_{1-2}(\boldsymbol{k})}\,, \tag{87b}$$

where

$$\chi^{(0)}_{1-2}(\boldsymbol{k}) = -T\sum_{34}\delta_{1-2,3-4}\,\varphi_{34}(\boldsymbol{k})\,. \tag{87c}$$

The above expressions provide all of the information necessary for one-loop renormalizations of two-point correlation functions in the density formalism.

V. A SIMPLE APPLICATION: PERTURBATION THEORY FOR THE DENSITY OF STATES

As a simple demonstration of how the formalism developed above works, let us consider the density of states (DOS). We first use the phase-space formalism from Sec. IV D. From Eq. (17) we see that the DOS is given by

$$\begin{aligned} N(\omega) &= N_{\rm F} + \mathrm{Re}\,Q(i\omega_n\to\omega+i0) \\ &\equiv N_{\rm F} + \delta N(\omega)\,, \end{aligned} \tag{88a}$$

where $N_{\rm F}$ is the DOS at the Fermi level including the interaction in Hartree-Fock approximation, and

$$Q(i\omega_n) = \frac{4}{\pi}\,\frac{1}{V}\sum_{\boldsymbol{p}}\left\langle {}^0_0 P_{nn}(\boldsymbol{p},\boldsymbol{p})\right\rangle\,. \tag{88b}$$

We next observe, from Eq. (63),

$$\langle P_{11}\rangle = \langle\Lambda_{11}\rangle + \Phi_{11}\,\frac{2\gamma T}{V}\sum_{2}\langle\Lambda_{22}\rangle\,. \tag{89}$$

The second contribution to $\langle P\rangle$ is proportional to the first one with an additional frequency integration which decreases the effects of the soft modes. The leading hydrodynamic contribution to Q is therefore given by Eq. (88b) with P replaced by Λ. Using Eq. (64) we can express Q in terms of a loop expansion in terms of the q-correlation functions. To one-loop order the expression for Λ in Eq. (64) suffices. Of the two contributions to $\langle\Lambda_{11}\rangle$, the second one has no hydrodynamic content due to the frequency restrictions inherent in Φ or $\hat{\gamma}q$. From the first contribution we obtain

$$Q(i\omega_{n_1}) = \frac{-2i}{\pi V}\sum_{\boldsymbol{p}_1}\sum_{3} G_3^{-1}\,\mathrm{tr}\,\left\langle \not{q}_{13}\,\not{q}^{+}_{31}\right\rangle\,, \tag{90}$$

where we neglect 2-loop contributions and less divergent terms. Inserting the Gaussian propagator, Eq. (68), and choosing $n_1 > 0$, we find

$$\begin{aligned} Q(i\omega_{n_1}) &= \frac{-2i\gamma}{\pi}\,\frac{1}{V}\sum_{\boldsymbol{k}} T\sum_{m<0}\varphi^{(3)\,++,-}_{n_1,n_1,m}(\boldsymbol{0},\boldsymbol{k}) \\ &\qquad\times\frac{1}{1+2\gamma\chi^{(0)}_{n_1-m}(\boldsymbol{k})}\,. \end{aligned} \tag{91}$$

Here $\varphi^{(3)\,++,-}$ is given by Eq. (81b), and we have explicitly indicated that the first two frequency arguments are positive while the third one is negative, and $\chi^{(0)}$ is given by Eq. (87c). The same result is obtained if we realize that δN is given by $\langle P_{nn}(\boldsymbol{p})\rangle$ in the density formalism of Sec. IV E and use Eqs. (84a), (85), and (87b).

To find the leading hydrodynamic contribution to Q or δN we can calculate $\varphi^{(3)}$ in the AGD approximation, whence it depends only on $n_1 - m$. Performing the integral at $T = \omega = 0$ in $d = 1 + \epsilon$ dimensions, we obtain

$$\delta N(\omega = 0)/N_{\rm F} = \frac{-1}{\epsilon}\,[1 + O(\epsilon)]\,\frac{G}{\pi}\,f(A_0^{\rm s}) + O(G^2)\ , \tag{92a}$$

where G is the loop expansion parameter from Eq. (71c),

$$f(x) = \frac{1 - x/2}{\sqrt{1-x}} - 1\ , \tag{92b}$$

and

$$A_0^{\rm s} = 2N_{\rm F}\gamma = \frac{2N_{\rm F}\Gamma}{1 + 2N_{\rm F}\Gamma} \tag{92c}$$

is the singlet s-wave Landau scattering amplitude in the particle-hole channel.

There are several interesting aspects of this result. The $1/\epsilon$ singularity reflects the instability of the Fermi-liquid against a Luttinger liquid in $d = 1$.[19,53–55] The current formalism emphasizes that this is a result of the Goldstone fluctuations getting so strong that they destroy the ordered phase. In fixed dimension $1 < d < 3$ as a function of the imaginary frequency the corresponding result is

$$(Q(i\omega_n) - Q(i0))/N_{\rm F} = c_d\ f_d(A_0^{\rm s})\,\omega_n^{d-1}\ , \tag{93a}$$

where $c_d > 0$ interpolates smoothly between $c_d \propto 1/(d-1)$ for $d \to 1$ and $c_d \propto 1/(3-d)$ for $d \to 3$. f_d is a function that continuously evolves to f given in Eq. (92b) as $d \to 1$, is positive definite for $1 < d \leq 3$, and has the property $f_d(x \to 0) \propto x^2$. For $d = 3$ one finds

$$\begin{aligned}(Q(i\omega_n) - Q(i0))/N_{\rm F}\ &=\ \tilde{c}_3\ f_3(A_0^{\rm s})\,\omega_n^2\log(1/\omega_n)\\ &\quad + O(\omega_n^2)\end{aligned} \tag{93b}$$

with $\tilde{c}_3 > 0$.

After analytically continuing to the real axis these results imply for the leading nonanalytic frequency correction to the density of states

$$\delta N(\omega)/N_{\rm F} = n_d\,(2-d)\,|\omega|^{d-1} \tag{94a}$$

for $1 < d < 3$, and

$$\delta N(\omega)/N_{\rm F} = -\tilde{n}_3\,\omega^2\log(1/|\omega|) \tag{94b}$$

for $d = 3$. n_d is positive definite and interpolates smoothly between $n_d \propto 1/(d-1)$ for $d \to 1$ and $n_d \propto 1/(3-d)$ for $d \to 3$, and $\tilde{n}_3 > 0$.

Note that to one-loop order there is no $|\omega|$ contribution to the DOS in $d = 2$, in agreement with earlier results obtained with diagrammatic techniques.[22,56–58] Equations (92a), (93), and (94) are analogous to the Coulomb or zero-bias anomaly in a disordered Fermi liquid.[14,59,60] There, the singularity at zero frequency is proportional to $1/(d-2)$, and the nonanalytic frequency dependence is $|\omega|^{(d-2)/2}$. The exponents of the respective frequency nonanalyticities reflect the scale dimensions of the least irrelevant operator at the disordered and clean Fermi-liquid fixed point, which are $-(d-2)$ and $-(d-1)$, and the corresponding dynamical exponents, which are $z = 2$ and $z = 1$, respectively.[16,19,54] If we denote the least irrelevant operator by u, this leads to a homogeneity law in the clean case

$$\delta N(\omega) = \delta N(\omega\, b, u\, b^{-(d-1)})\ , \tag{95}$$

which immediately leads to the frequency dependence reflected in Eq. (94a). The sign of the effect at zero frequency, Eq. (92a), reflects the fact that a repulsive interaction will lead to a decrease of the density of states at the Fermi level.

The dependence of the anomaly on the interaction amplitude also warrants a comment. The function $f(x)$, Eq. (92b), is positive definite for $0 < x < 1$, and $f(x \to 0) \propto x^2$. The effects of the soft modes thus appear only at second order in the interaction, whereas Eq. (90) or (91) naively suggest an effect at first order. The technical reason is that the integral in Eq. (91) vanishes if one puts $\gamma = 0$ in the denominator of the integrand. This is consistent with many-body perturbation theory, as it must be: The diagram in Fig. 2 (a), which is the only contribution to the DOS to first order in the interaction (Hartree diagrams have been absorbed into the Green function), does not allow for a

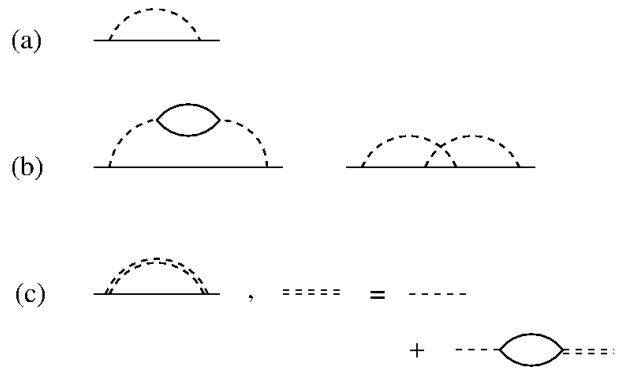

FIG. 2: Diagrammatic contributions to the DOS at (a) first and (b) second order in the interaction amplitude. (c) Shows a ladder resummation of the diagram in (a) plus the first diagram in (b), which is identical with Eq. (91).

mixing of retarded and advanced degrees of freedom and hence has no hydrodynamic content. The first diagram in Fig. 2 (b) does, and its hydrodynamic part is identical with Eq. (91) to second order in γ. The hydrodynamic part of its ladder resummation, Fig. 2 (c), is identical with the full Eq. (91). The other contribution to second order in the interaction, i.e., the second diagram in Fig. 2 (b), represents a particle-particle channel contribution. In the present formalism it will appear if a particle-particle interaction amplitude is added to Eqs. (4b) or (13).[42] Note that the effective theory picks out only the hydrodynamic parts of these diagrams, whereas the many-body diagrams contain non-hydrodynamic pieces as well. Also note that the current theory, at any given order in the loop expansion, in valid to all orders in the interaction amplitude, whereas many-body perturbation theory by necessity involves an expansion in powers of the interaction.

In addition to the above discussion there are other interesting aspects of the DOS that include the behavior in the case of a long-ranged Coulomb interaction,[56,61] as well as the two-point correlation function of the local DOS. These effects will be discussed elsewhere.[42]

VI. SUMMARY, DISCUSSION, AND OUTLOOK

A. Summary and Discussion

We have derived an effective field theory that describes the soft modes, and their effects on observables, in electron systems without quenched disorder. The soft modes have been identified by means of a Ward idendity; they are particle-hole excitations with a linear frequency-momentum relation. They are the Goldstone modes of a continuous symmetry, namely, a rotation in Matsubara frequency space that relates retarded and advanced degrees of freedom and is spontaneously broken whenever the quasiparticle spectral weight is nonzero. This identifies a Fermi liquid as an ordered phase with the quasiparticle spectral weight as the order parameter and the soft particle-hole excitations the corresponding Goldstone modes. A nonzero temperature the Goldstone modes acquire a mass, and so do subsets of them in the presence of a magnetic field or other symmetry-breaking external fields. Quenched disorder restricts the Goldstone modes to the s-wave or $\ell = 0$ channel, whereas all higher angular momentum channels become massive. This is the reason why the clean theory is more complicated than the corresponding one for disordered electrons: There are many more soft modes to take into account in clean systems. It is important to remember that the Goldstone modes are *not* density excitations, and their existence is not related to particle-number conservation; the density susceptibility is characterized by zero-sound excitations regardless of whether the symmetry is broken or not, and at any temperature.

Technically, we have borrowed heavily from techniques and concepts developed by Wegner[8] for disordered noninteracting electrons and generalized by Finkelstein[12] to the interacting disordered case. However, the resulting effective field theory in the clean case does *not* take the form of a generalized nonlinear sigma model, as it does in the disordered case, but is significantly more complex. This complexity reflects the presence of many more soft modes. One manifestation of this complexity is the fact that the effective action contains terms to higher than linear order in the interaction amplitude, whereas in the disordered case only linear terms appear. The reason is that in the clean case the Green function is soft, and the additional powers of the frequency that accompany higher powers of the interaction are compensated by multiplicative Green functions. In the disordered case, in contrast, the Green function is massive because of the elastic relaxation rate, and terms of higher than linear order in the interaction are irrelevant in a

renormalization-group sense.

We have followed Wegner[8] and later work based on his ideas[9–11,14,16] in bosonizing the theory by introducing classical matrix fields that are isomorphic to bilinear products of fermion fields. Technically, this is done by introducing a Lagrange multiplier constraint and integrating out the fermions. This approach is very different from that of Shankar,[24] who applied renormalization-group techniques directly to the fermionic theory to derive Landau Fermi-liquid theory. The Lagrange multiplier field, whose expectation value plays the role of a self energy, is soft and needs to be kept to all orders to avoid introducing spurious soft modes. This has been achieved in terms of diagram rules for the effective theory, with the role of the Lagrange multiplier being to eliminate the noninteracting part of certain propagators. The effective action has not been given in closed form, but an explicit method has been devised to construct it to any desired power in the fundamental soft-mode field. This allows for a systematic loop expansion, with the inverse Fermi velocity playing the role of the bare loop expansion parameter. The interaction, by contrast, is *not* treated perturbatively. Order by order in a loop expansion, the theory thus produces the leading hydrodynamic (i.e., long-wavelength and low-frequency) effects to all orders in the interaction, and the interaction is not required to be weak in any sense. This, together with its ability to isolate the hydrodynamic contributions to any desired observable, sets the current theory apart from many-body perturbation theory which expands in powers of the interaction, with any resummations to all orders in the interaction being opportunistic and not controlled.

A related feature of the current theory that is even more important is its amenability to renormalization-group techniques. Since it is formulated in terms of an infinite power series in the fundamental field, with each term characterized by an n-point vertex and a corresponding correlation function, the structure of which is dictated by symmetry, the bare theory also suggests the structure of the corresponding renormalized theory, which many-body perturbation theory cannot do. This feature has been exploited extensively in the disordered case to describe, inter alia, the Anderson-Mott metal-insulator transition,[13,14,62] which had proven very hard to do by means of many-body theory.

Certain aspects of the present theory had already been present in the theory put forward in Ref. 16. However, since the latter focused on disordered systems, it was formulated in terms of s-wave variables, with the higher angular momentum modes effectively being integrated out. In an application to the clean case, this led to a nonlocal theory, i.e., the 3-point and higher vertices did not exist in the hydrodynamic limit of small frequencies and wave numbers. In addition, no complete separation of soft and massive modes was achieved in that theory for the clean case. Both of these points have been remedied in the present formulation. The issue of the higher-angular-momentum modes is handled by formulating the theory in terms of phase-space variables, which explicitly keeps all angular momentum channels, and the separation of soft and massive modes has been achieved by expressing the massive modes in terms of the soft ones in a power-series expansion, albeit not in a closed form, in contrast to the disordered case. The relation to the formulation in terms of s-wave variables, which is still useful for certain calculational purposes, has been explained in Sec. IV E.

B. Outlook

Formalism for formalism's sake is not very useful, and a crucial test for the present theory will be any physical problems it can solve more easily than other methods. We conclude this paper by listing several applications that we plan to pursue in the near future, many of which we already have obtained preliminary results for.

The density-of-states calculation in Sec. V is meant only as an illustration of how the loop-expansion-based perturbation theory works within the present context; it does not achieve anything that cannot easily be done with many-body diagrams. It nevertheless has some important advantages, and points to some crucial future developments. For instance, as was shown in Ref. 16, the Fermi-liquid state is represented by a stable renormalization-group fixed point, and the leading irrelevant operators with respect to that fixed point can be easily identified. This shows that the nonanalytic energy or frequency dependence of the DOS near the Fermi level, Eqs. (93), is qualitatively exact; there cannot be any contributions with a lower power of the frequency. This would be impossible to establish within many-body theory. A related point is that the field theory, in contrast to many-body perturbation theory, is not perturbative in the interaction; the results at any order in the loop expansion are valid to all orders in the interaction. Other observables will also display universal nonanalytic behavior in the Fermi-liquid phase that can be understood as corrections to scaling at the Fermi-liquid fixed point, and they can be identified by the same techniques. One interesting quantity is the two-point correlation function of the local DOS. Since the DOS is the (zeroth moment of) the order parameter for the Fermi liquid, this plays the role of the order parameter susceptibility, and it is expected to be long-ranged, i.e., to exhibit a universal divergence for small wave numbers, everywhere in the Fermi-liquid phase. This expected phenomenon is analogous to a well-known feature of the magnetic susceptibility in the ordered phase of a Heisenberg ferromagnet.[6] In Ref. 30 we have used scaling arguments combined with a preliminary version of the current theory to predict that the DOS-DOS correlation function diverges as T/k^{3-d} for small wave numbers k. We

will revisit this problem, together with other aspects of the DOS, in more detail shortly.

In addition, the effective theory suggests the existance of a critical fixed point that describes a transition from a Fermi liquid to a non-Fermi-liquid state. This will be analogous in many respects to the critical fixed points that describe various versions of the Anderson-Mott transition.[14] The basic physical idea is that the fluctuations in a Fermi liquid that are represented by the Goldstone modes become stronger with decreasing dimensionality and increasing interaction strength, and ultimately give rise to an instability of the Fermi liquid since the system can lower its free energy by going into a state that restores the symmetry and thus eliminates the Goldstone modes. A scaling theory for such a transition has been developed in Ref. 30, and arguments have been given that suggest that the corresponding fixed point will appear in the present theory at 2-loop order. The loop expansion will then allow for a controlled description of the transition in $d = 1 + \epsilon$ dimensions.

The theory can also be applied to magnetic phenomena, i.e., order in the spin-triplet channel. In Sec. IV A 2 we have constructed a saddle-point solution that reproduces the Stoner theory of ferromagnetism. By expanding about that saddle point, analogously to the expansion about the Fermi-liquid saddle point in Sec. IV B, one can construct a systematic theory for quantum Heisenberg ferromagnets that improves upon and replaces the Hertz-Millis theory.[63,64] Such a theory has been developed in Ref. 65, and it has been shown to have very interesting consequences, including the fact that the phase transition at sufficiently low temperatures in a Heisenberg ferromagnet is generically of first order, in agreement with experimental observations. This theory was based on a phenomenological treatment of the fermionic degrees of freedom combined with symmetry arguments, and it does not have the correct limiting behavior as $d \to 1$. The current theory will allow for a systematic derivation valid in any dimension $d > 1$ that will serve as an important check and will probably point to new developments. It can also be applied to ferrimagnets, the quantum aspects of which have not been considered so far. Finally, if one considers band electrons instead of the nearly-free electron model we have employed for simplicity, the theory can also describe antiferromagnetism.

Treatments of the spin-triplet channel are not restricted to s-wave order. By keeping a triplet interaction amplitude in the p-wave channel, see Eq. (13d), one can describe a magnetic nematic or p-wave ferromagnet. In Ref. 29 we have shown that the quantum phase transition from a Fermi liquid to such magnetic state is generically of first order, for reasons that are similar to those for the analogous statement for the s-wave ferromagnetic transition. This is in contrast to Hertz-type theories[28] that predict a continuous transition. This result was also based on a preliminary version of the current theory, and the present complete formulation will allow for more detailed studies of exotic magnetic states. The same is true for superconducting states, in either the s-wave or higher angular momentum channels.

All of the above examples involve phenomena that crucially rely on the coupling of some order parameter to fermionic soft modes. The present effective theory provides a general framework for describing the universal aspects of ordered phases, and the quantum phase transitions between them, of clean electron systems.

Acknowledgments

We gratefully acknowledge helpful correspondence with Andrey Chubukov and Dmitrii Maslov regarding the density of states in 2-d systems. This work was supported by the National Science Foundation under Grant Nos. DMR-09-29966, and DMR-09-01907.

Appendix A: Effects of quenched disorder

In this appendix we show that quenched disorder modifies the Ward identity we discussed in Sec. III such that only density or $\ell = 0$ modes remain soft.

Consider a static random potential $u(\boldsymbol{x})$ that couples to the electron number density and has a Gaussian distribution with a second moment

$$\{u(\boldsymbol{x})\, u(\boldsymbol{y})\}_{\rm dis} = U(\boldsymbol{x} - \boldsymbol{y})/\pi N_{\rm F} \ , \tag{A1}$$

where $\{\ldots\}_{\rm dis}$ denotes the disorder average and U is dimensionally an energy density. We then need to add to Eq.

(13a) a contribution

$$\begin{aligned}\mathcal{A}_{\rm dis} &= \frac{-1}{2\pi N_{\rm F}}{\sum_{\boldsymbol{q}}}' U(\boldsymbol{q}) \sum_{\boldsymbol{k}} {\rm tr}\ (Q(\boldsymbol{k};\boldsymbol{q}))\ {\rm tr}\ (Q(\boldsymbol{p};-\boldsymbol{q})) \\ &+\frac{1}{\pi N_{\rm F}}{\sum_{\boldsymbol{q}}}' \sum_{\boldsymbol{k},\boldsymbol{p}} U(\boldsymbol{k}-\boldsymbol{p})\, {\rm tr}\ (Q(\boldsymbol{k};\boldsymbol{q})\, Q(\boldsymbol{p};-\boldsymbol{q}))\ .\end{aligned} \tag{A2}$$

The trace now includes tracing over replica indices to handle the disorder average. Now we anticipate that the center-of-mass momenta $\boldsymbol{k}$ and $\boldsymbol{p}$ will be pinned to the Fermi surface, and expand $U(\boldsymbol{k}-\boldsymbol{p})$ in Legendre polynomials. We then can write the disorder part of the action

$$\begin{aligned}\mathcal{A}_{\rm dis} &= \frac{-1}{2\pi N_{\rm F}\tau_1}{\sum_{\boldsymbol{q}}}' {\rm tr}\ \left(Q^{0,0}(\boldsymbol{q})\right)\ {\rm tr}\ \left(Q^{0,0}(-\boldsymbol{q})\right) \\ &+\frac{1}{\pi N_{\rm F}} \sum_{\ell=0}^{\infty} \frac{1}{\tau_{\rm rel}^{(\ell)}} \frac{1}{k_{\rm F}^{2\ell}} \sum_{m=-\ell}^{\ell} (-)^m {\sum_{\boldsymbol{q}}}' \\ &\times {\rm tr}\ \left(Q^{\ell,m}(\boldsymbol{q})\, Q^{\ell,-m}(-\boldsymbol{q})\right)\ .\end{aligned} \tag{A3a}$$

Here we have expanded the Q-matrices in spherical harmonics Y_ℓ^m,

$$Q_{12}^{\ell,m}(\boldsymbol{q}) = \sum_{\boldsymbol{k}} k^\ell\, Y_\ell^m(\Omega_{\boldsymbol{k}})\, Q_{12}(\boldsymbol{k};\boldsymbol{q})\ , \tag{A3b}$$

and have defined relaxation rates $1/\tau_1 = U(\boldsymbol{q}=0)$ and

$$\frac{1}{\tau_{\rm rel}^{(\ell)}} = 2\pi \int_{-1}^{1} d\eta\, P_\ell(\eta)\, U(k_{\rm F}\sqrt{2(1-\eta)}) \tag{A3c}$$

in the angular momentum channels.

The derivation of the Ward identity as in Sec. III now still yields Eq. (30) (note $Q^{0,0} \equiv Q^{(0)}$), but it gives no information about the angular momentum channels with $\ell > 1$. This is consistent with the results of perturbation theory. The Fermi-liquid saddle-point solution from Sec. IV A 1 needs to be augmented by a disorder-induced self energy and becomes identical with the saddle point considered in Ref. 16. An expansion about the saddle point to Gaussian order yields

$$\begin{aligned}&\langle \delta Q_{12}^{\ell,m}(\boldsymbol{q}=0)\, Q_{12}^{\ell,m}(\boldsymbol{q}=0)\rangle_{i\Omega_{n_1-n_2}\to 0} \\ &\qquad \propto \frac{\pi N_{\rm F}}{1/\tau_{\rm rel}^{(\ell=0)} - 1/\tau_{\rm rel}^{(\ell)}}\ .\end{aligned} \tag{A4}$$

In order for the theory to be stable we need to require $1/\tau_{\rm rel}^{(\ell=0)} > 1/\tau_{\rm rel}^{(\ell>0)}$, which will generically be the case. If it is not one needs to choose a saddle point in the angular-momentum channel that has the largest value of $1/\tau_{\rm rel}^{(\ell=0)}$. We see that all correlation functions other than the one for $\ell = 0$ are massive, which is consistent with the information obtained from the Ward identity.

Appendix B: An illustrative example: Soft modes in the classical $O(2)$ vector model

In this appendix we discuss a classical $O(2)$ vector model as an illustrative example of how to separate soft and massive modes. Consider a 2-component classical field $\boldsymbol{\phi}(\boldsymbol{x}) = (\pi(\boldsymbol{x}), \sigma(\boldsymbol{x}))$ and an action

$$\mathcal{A} = \int d\boldsymbol{x}\ \left[\frac{r}{2}\,\boldsymbol{\phi}^2(\boldsymbol{x}) + \frac{c}{2}\,\left(\nabla\boldsymbol{\phi}(\boldsymbol{x})\right)^2 + \frac{u}{4}\,\left(\boldsymbol{\phi}^2(\boldsymbol{x})\right)^2\right] \tag{B1}$$

with $(\nabla\boldsymbol{\phi})^2 = \partial_i\phi_j\partial^i\phi^j$. The partition function is

$$Z = \int D[\boldsymbol{\phi}]\ e^{-\mathcal{A}[\boldsymbol{\phi}]}\ . \tag{B2}$$

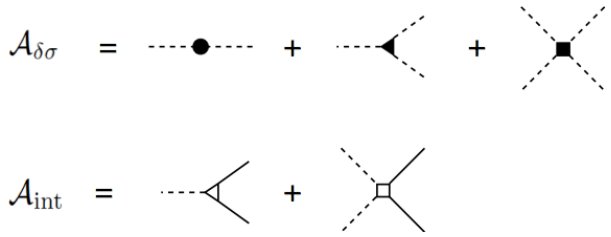

FIG. 3: Graphic representation of the contributions $\mathcal{A}_{\delta\sigma}$ and $\mathcal{A}_{\rm int}$ to the action. Dashed and solid lines represent the fields $\delta\sigma$ and π, respectively.

The saddle-point equations for homogeneous field configurations (π, σ) read

$$r\pi + u\pi^3 + u\pi\sigma^2 \;=\; 0 \;, \tag{B3a}$$
$$r\sigma + u\sigma^3 + u\sigma\pi^2 \;=\; 0 \;. \tag{B3b}$$

For $r < 0$, a solution that minimizes the free energy is

$$\pi_{\rm sp} \;=\; 0 \;, \tag{B4a}$$
$$\sigma_{\rm sp} \;=\; \sqrt{-r/u} \equiv \phi_0 \;. \tag{B4b}$$

Alternatively, we can choose $\pi_{\rm sp}$ to have any value with $|\pi_{\rm sp}| < \phi_0$. Then

$$\sigma_{\rm sp}^2 = \phi_0^2(1 - \pi_{\rm sp}^2/\phi_0^2) \tag{B5}$$

also solves the saddle-point equations. This reflects the $O(2)$ rotational symmetry of the model.

Now consider fluctuations about the saddle point, $\sigma = \phi_0(1+\delta\sigma)$, and redefine $\pi \to \phi_0\pi$, so that $\delta\sigma$ and π are both dimensionless. Then the action takes the form

$$\mathcal{A} = \mathcal{A}_{\rm sp} + \mathcal{A}_\pi[\pi] + \mathcal{A}_{\delta\sigma}[\delta\sigma] + \mathcal{A}_{\rm int}[\pi, \delta\sigma] \;, \tag{B6a}$$

where

$$\mathcal{A}_\pi[\pi] \;=\; \frac{c}{2}\,\phi_0^2 \int d\boldsymbol{x}\;\; (\boldsymbol{\nabla}\pi)^2 + \frac{u}{4}\,\phi_0^4 \int d\boldsymbol{x}\;\, \pi^4 \;, \tag{B6b}$$

$$\begin{aligned}\mathcal{A}_{\delta\sigma}[\delta\sigma] \;=\; & u\phi_0^4 \int d\boldsymbol{x}\;\; (\delta\sigma)^2 + \frac{c}{2}\,\phi_0^2 \int d\boldsymbol{x}\;\; (\boldsymbol{\nabla}\delta\sigma)^2 \\ & + u\phi_0^4 \int d\boldsymbol{x}\;\, (\delta\sigma)^3 + \frac{u}{4}\,\phi_0^4 \int d\boldsymbol{x}\;\, (\delta\sigma)^4 \;,\end{aligned} \tag{B6c}$$

$$\mathcal{A}_{\rm int}[\pi, \delta\sigma] \;=\; u\phi_0^4 \int d\boldsymbol{x}\;\, \pi^2\,\delta\sigma + \frac{u}{2}\,\phi_0^4 \int d\boldsymbol{x}\;\, \pi^2(\delta\sigma)^2 \;. \tag{B6d}$$

Here $\delta\sigma$ and π are functions of $\boldsymbol{x}$. A graphic representation of the terms that depend on $\delta\sigma$ is shown in Fig. 3.

It is well known that the transverse fluctuation, i.e. the field π, is the soft mode in this model.[7] π is the Goldstone mode that reflects the spontaneously broken $O(2)$ invariance, and it is governed by a Ward identity that ensures that the homogeneous transverse susceptibility diverges. The effective soft-mode theory is a nonlinear sigma model with an action

$$\mathcal{A}_{\rm NL\sigma M} = \frac{c}{2}\,\phi_0^2 \int d\boldsymbol{x}\;\, \left[(\boldsymbol{\nabla}\pi)^2 + (\boldsymbol{\nabla}\sqrt{1-\pi^2})^2\right] \tag{B7a}$$

and a partition function

$$Z = \int \frac{D[\pi]}{\sqrt{1-\pi^2}}\, e^{-\mathcal{A}_{\rm NL\sigma M}[\pi]} \;. \tag{B7b}$$

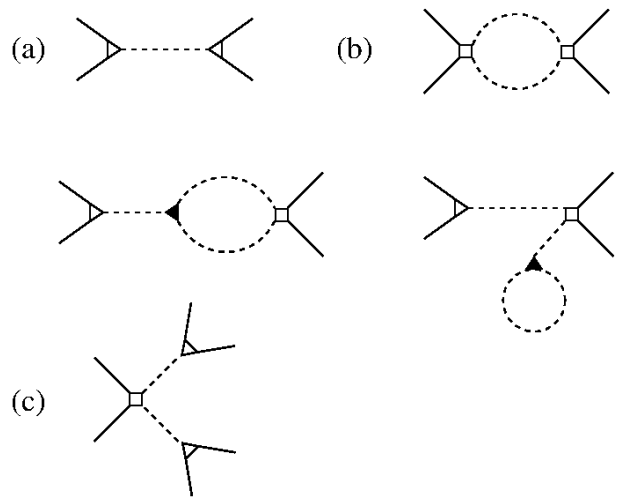

FIG. 4: (a) Tree-level and (b) one-loop contributions to the term of $O(\pi^4)$ in $\Delta\mathcal{A}_\pi[\pi]$ in an expansion in $\delta\sigma$ loops. (c) Tree-level contribution to the term of $O(\pi^6)$.

Notice that the functional integration measure has changed by switching to π as the only field. This is important to preserve the symmetry of the problem. Technically, terms arising from the measure cancel spurious mass terms that get generated in perturbation theory.

In this simple case the nonlinear sigma model can be derived in closed form by writing $\boldsymbol{\phi}$ as a representation of $O(2)$ rotations,

$$\boldsymbol{\phi}(\boldsymbol{x}) = \begin{pmatrix} \sqrt{1-\pi^2(\boldsymbol{x})} & \pi(\boldsymbol{x}) \\ -\pi(\boldsymbol{x}) & \sqrt{1-\pi^2(\boldsymbol{x})} \end{pmatrix} \begin{pmatrix} 0 \\ \rho(\boldsymbol{x}) \end{pmatrix} , \tag{B8}$$

and integrating out the massive ρ-fluctuations in saddle-point approximation (i.e., replacing $\rho(\boldsymbol{x})$ by its saddle-point value ϕ_0, see Ref. 7). For the more complicated matrix field theory we are interested in no closed-form derivation of the effective soft-mode theory has been found so far, but one can still derive the latter perturbatively order by order in powers of the soft modes. To guide that derivation it is useful to consider the analogous process for the $O(2)$ model. In what follows, we thus pretend that the nonlinear sigma model were not known, and construct an effective soft-mode theory perturbatively by considering Eqs. (B6) and systematically integrating out $\delta\sigma$.

Returning to Eqs. (B6) we see that at the Gaussian level π is indeed massless and $\delta\sigma$ is massive. However, the π^4 term in Eq. (B6b) does not carry any gradients. If we were to simply neglect $\delta\sigma$ this term would generate a mass for π at the one-loop level. This mass violates the Ward identity and is spurious; it gets canceled if one properly integrates out $\delta\sigma$. The coupling between $\delta\sigma$ and π is thus crucial and needs to be kept. Formally integrating out $\delta\sigma$ leads to an effective action in terms of π only,

$$\mathcal{A}_{\rm eff}[\pi] = \mathcal{A}_\pi[\pi] + \Delta\mathcal{A}_\pi[\pi] \tag{B9a}$$

where

$$\Delta\mathcal{A}_\pi[\pi] = -\ln \int D[\delta\sigma]\; e^{-\mathcal{A}_{\delta\sigma}[\delta\sigma]-\mathcal{A}_{\rm int}[\pi,\delta\sigma]} . \tag{B9b}$$

We can now perform the $\delta\sigma$-integral perturbatively.

Let us first integrate out $\delta\sigma$ in a straightforward manner. The resulting loop expansion is an expansion in powers of $1/\phi_0$. The tree-level and one-loop diagrams that contribute to $O(\pi^4)$ are shown in Fig. 4. The tree diagram yields

$$\Delta\mathcal{A}_\pi[\pi] \approx -\frac{u}{4}\,\phi_0^4 \int d\boldsymbol{x}\; \pi^4 + \frac{c}{8}\,\phi_0^2 \int d\boldsymbol{x}\; (\boldsymbol{\nabla}\pi^2)^2 \tag{B10}$$

Combining Eq. (B10) with Eq. (B6b) we see that the offending π^4 term cancels, and the gradient terms agree with the expansion of the nonlinear sigma model, Eq. (B7a), to order π^4. Terms of higher order in π can be constructed analogously, see Fig. 4(c).

Alternatively, we can evaluate the integral in Eq. (B9b) by means of a saddle-point approximation for $\delta\sigma$ for fixed π. The saddle-point equation is shown graphically in Fig. 5. Analytically, we obtain

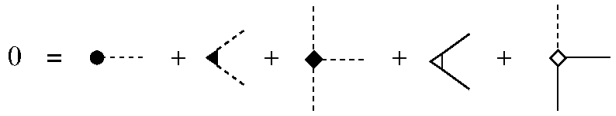

FIG. 5: Saddle-point equation for $\delta\sigma$ for fixed π.

$$0 \;=\; \frac{1}{2}\,\pi^2(\boldsymbol{x}) + \left(1+\frac{1}{2}\,\pi^2(\boldsymbol{x})\right)\delta\sigma_{\rm sp}(\boldsymbol{x}) \\ -\frac{c}{2u\phi_0^2}\,\boldsymbol{\nabla}^2\delta\sigma_{\rm sp}(\boldsymbol{x}) + \frac{3}{2}\,\left(\delta\sigma_{\rm sp}(\boldsymbol{x})\right)^2 + \frac{1}{2}\,\left(\delta\sigma_{\rm sp}(\boldsymbol{x})\right)^3 . \tag{B11a}$$

In a gradient expansion, and keeping terms up to $O(\nabla^2)$, this is solved by

$$\delta\sigma_{\rm sp}(\boldsymbol{x}) \;=\; \sqrt{1-\pi^2(\boldsymbol{x})} - 1 \\ +\frac{g}{1-\pi^2(\boldsymbol{x})}\,\boldsymbol{\nabla}^2\sqrt{1-\pi^2(\boldsymbol{x})} + O(g^2\boldsymbol{\nabla}^4)\ , \tag{B11b}$$

with $g = c/2u\phi_0^2$. Note that higher powers of g are accompanied by higher powers of gradients. Substituting this result back into the action, Eqs. (B6), we obtain the action of the nonlinear sigma model, Eq. (B7a).

In contrast to the $\delta\sigma$-loop expansion above, this second method does not require knowledge of the $\delta\sigma$ propagator. It also involves only solving an algebraic equation, no integrals need to be performed. In contrast to the standard derivation of the nonlinear sigma model, which uses Eq. (B8), it reqires no explicit knowledge of the symmetry group and its representations. If no closed-form solution of the saddle-point equation could be found, one could still solve Eq. (B11a) iteratively, since it is inhomogeneous. In this way one obviously can construct the effective action perturbatively, order by order in powers of π. For an analysis of the effective action to any given order in a loop expansion this would be sufficient. This is the procedure we follow for the matrix field theory in Sec. IV C.

Finally, we need to worry about the change of the integration measure mentioned after Eqs. (B7). In the present formalism, this is generated by the Gaussian fluctuations about $\delta\sigma_{\rm sp}$. If we write $\delta\sigma(\boldsymbol{x}) = \delta\sigma_{\rm sp}(\boldsymbol{x}) + \tau(\boldsymbol{x})$ and integrate out $\tau(\boldsymbol{x})$ in Gaussian approximation we obtain an additional contribution to the effective action. The complete effective action then is

$$\mathcal{A}_{\rm eff} \;=\; \frac{1}{2G}\int d\boldsymbol{x}\,\left[(\boldsymbol{\nabla}\pi)^2 + (\boldsymbol{\nabla}\sqrt{1-\pi^2})^2\right] \\ +{\rm Tr}\,\ln\sqrt{1-\pi^2} + O(G), \tag{B12}$$

with $G = 1/c\phi_0^2$ the coupling constant. This is the nonlinear sigma model including the measure terms, see Eqs. (B7).

[1] E. M. Lifshitz and L. P. Pitaevskii, *Statistical Physics, Part 2* (Pergamon, New York, 1980).
[2] G. Baym and C. Pethick, *Landau Fermi-Liquid Theory* (Wiley, New York, 1991).
[3] A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Dover, New York, 1963).
[4] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, New York, 1971).
[5] S.-K. Ma, *Modern Theory of Critical Phenomena* (Benjamin, Reading, MA, 1976).
[6] E. Brézin and D. J. Wallace, Phys. Rev. B **7**, 1967 (1973).
[7] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena* (Oxford University Press, Oxford, 1996).
[8] F. Wegner, Z. Phys. B **35**, 207 (1979).
[9] L. Schäfer and F. Wegner, Z. Phys. B **38**, 113 (1980).
[10] A. M. M. Pruisken and L. Schäfer, Nucl. Phys. B **200** [**FS4**], 20 (1982).
[11] K. B. Efetov, A. I. Larkin, and D. E. Khmelnitskii, Zh. Eksp. Teor. Fiz. **79**, 1120 (1980), [Sov. Phys. JETP **52**, 568 (1980)].
[12] A. M. Finkelstein, Zh. Eksp. Teor. Fiz. **84**, 168 (1983), [Sov. Phys. JETP **57**, 97 (1983)].
[13] A. M. Finkelstein, Zh. Eksp. Teor. Fiz. **86**, 367 (1984), [Sov. Phys. JETP **59**, 212 (1984)].
[14] D. Belitz and T. R. Kirkpatrick, Rev. Mod. Phys. **66**, 261 (1994).

[15] The effective theories for disordered electrons quoted above all ignore effects that are nonperturbative with respect to the disorder, such as Griffiths singularities; see, e.g., Ref. 66 and references therein. Statements to the effect that our effective theory for clean systems is more complicated are always meant with respect to the existing theories for disordered electrons that are incomplete in this sense. A complete theory that takes such effects into account might well be much more complicated than either the existing theories for disordered systems or the theory for clean ones presented in this paper.

[16] D. Belitz and T. R. Kirkpatrick, Phys. Rev. B **56**, 6513 (1997).

[17] I. L. Aleiner and K. B. Efetov, Phys. Ref. B **74**, 075102 (2006).

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

[19] W. Metzner, C. Castellani, and C. D. Castro, Adv. Phys. **47**, 317 (1998).

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

[21] A. Chubukov and D. Maslov, Phys. Rev. B **69**, 121102 (2004).

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

[23] A. Chubukov and D. Maslov, Phys. Rev. B **76**, 165111 (2007).

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

[25] K. B. Efetov, C. P'epin, and H. Meier, Phys. Rev. Lett. **103**, 186403 (2009).

[26] K. B. Efetov, C. P'epin, and H. Meier, Phys. Rev. B **82**, 235120 (2010).

[27] H. Meier, C. P'epin, and K. B. Efetov, Phys. Rev. B **84**, 205131 (2011).

[28] V. Oganesyan, S. A. Kivelson, and E. Fradkin, Phys. Rev. B **64**, 195109 (2001).

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

[30] T. R. Kirkpatrick and D. Belitz, arXiv:1108.2811.

[31] J. W. Negele and H. Orland, *Quantum Many-Particle Systems* (Addison-Wesley, New York, 1988).

[32] This statement holds in a nonrelativistic theory where the underlying microscopic equation of motion is the Schrödinger equation. Taking into account relativistic corrections to the Schrödinger equation leads, *inter alia*, to the dipole-dipole interaction, see Ref. 34. In what follows we will neglect all relativistic corrections.

[33] See Sec. 18 in Ref. 3.

[34] D. Belitz and T. R. Kirkpatrick, Phys. Rev. B **81**, 184419 (2010).

[35] The such defined adjoint has the properties $(\eta^+)^+ = \eta$, and $(z\eta)^+ = z\eta^+$ for any bispinor η and any complex number z. It is unrelated to the adjoint operation for the fermion fields, which obeys $\overline{\bar{\psi}} = \psi$ and $\overline{z\psi} = z^*\bar{\psi}$. Our notation is the same as in Ref. 16.

[36] S. V. Maleev and B. P. Toperverg, Zh. Eksp. Teor. Fiz. **69**, 1440 (1975), [Sov. Phys. JETP **42**, 734 (1976)].

[37] D. Vollhardt and P. Wölfle, Phys. Rev. B **22**, 4666 (1980).

[38] A. J. McKane and M. Stone, Ann. Phys. (NY) **131**, 36 (1981).

[39] T. Kirkpatrick and D. Belitz, Phys. Rev. B **65**, 195123 (2002).

[40] T. R. Kirkpatrick and D. Belitz, Phys. Rev. B **62**, 952 (2000).

[41] T. Moriya, *Spin Fluctuations in Itinerant Electron Magnetism* (Springer, Berlin, 1985).

[42] D. Belitz and T. R. Kirkpatrick, unpublished results.

[43] Note that the contour for the $\tilde{\Lambda}$ integration must start at $-i\infty$ and end at $i\infty$, as can be seen from Eq. (11). The sign of the $\bar{\Lambda}$ propagator is thus consistent with a stable field theory.

[44] We note again that the soft modes in the particle-particle channel and two of the particle-hole spin-triplet channels can be made massive by applying a magnetic field. This mass-generation mechanism is the same as in the disordered case.[14] Nevertheless, the spin-triplet and particle-particle degrees of freedom produce important physical effects, e.g., the mass renormalization in Fermi-liquid theory.[3] The current theory including the spin-triplet and particle-particle degrees of freedom will be discussed elsewhere.[42].

[45] Here and in what follows we denote by $O(q^n)$ terms that are of n^{th} order in q, λ, or any combination thereof.

[46] This is the same observation that was made in Ref. 16 in the context of a theory expressed in terms of $\delta\bar{Q}$ and $\delta\bar{\Lambda}$. The crucial difference between this treatment and the current one is that Ref. 16 considered a formulation in terms of density fluctuations. Soft modes in higher angular-momentum channels had therefore implicitly been integrated out at an early stage, and it was no longer obvious how to separate the soft modes from the massive ones.

[47] Note that, for instance, the first term in Eq. (67c) can *not* be simplified to read $G(\hat{\gamma}q + \hat{\gamma}q^+)G(\not{q}G^{-1}\not{q}^+ + \not{q}^+G^{-1}\not{q})$, lest the frequency restrictions inherent in Φ get lost.

[48] Note that this effective theory cannot be used to calculate q-vertex functions without imposing additional modifications of the usual diagram rules. The reason is that there are diagrams that are reducible with respect to λ, but irreducible with respect to q, and hence contribute to q-vertices, that are missed if one simply drops λ from the effective action. See Sec. IV.A of Ref. 16, where this point was already noted.

[49] F. Stillinger, J. Math. Phys. **18**, 1224 (1977).

[50] X.-F. He, Phys. Rev. B **43**, 2063 (1991).

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

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

[53] J. Sólyom, Adv. Phys. **28**, 201 (1979).

[54] C. Castellani, C. D. Castro, and W. Metzner, Phys. Rev. Lett. **72**, 316 (1998).

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

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

[57] We thank Andrey Chubukov and Dmitrii Maslov for pointing this out.

[58] This holds for a point-like interaction. In the case of a long-ranged Coulomb interaction the behavior in $d = 2$ is different, see Refs. 56,61.
[59] B. L. Altshuler and A. G. Aronov, Solid State Commun. **30**, 115 (1979).
[60] P. A. Lee and T. V. Ramakrishnan, Rev. Mod. Phys. **57**, 287 (1985).
[61] D. V. Khveshchenko and M. Reizer, Phys. Rev. B **57**, R4245 (1998).
[62] A. M. Finkelstein, Z. Phys. B **56**, 189 (1984).
[63] J. Hertz, Phys. Rev. B **14**, 1165 (1976).
[64] A. J. Millis, Phys. Rev. B **48**, 7183 (1993).
[65] T. R. Kirkpatrick and D. Belitz, Phys. Rev. B **67**, 024419 (2003).
[66] T. Vojta, Phys. Rev. Lett. **90**, 107202 (2003).